Department of Statistics PhD Thesis

2015

Prediction of Final Stature from a New 3D Two-Stage Growth Model

Shahin, Md. Abu
University of Rajshahi

Prediction of Final Stature from a New 3D Two–Stage Growth Model

A Dissertation Submitted to the Department of Statistics University of Rajshahi, Bangladesh in Partial Fulfillments of the Requirements for the Degree of Doctor of Philosophy

SUPERVISED BY

Dr. Md. Ayub Ali

Professor

Department of Statistics

University of Rajshahi

Rajshahi-6205

Bangladesh

SUBMITTED BY

Md. Abu Shahin
Ph.D. Fellow
Roll No.: 12228
Registration No.: 0567
Session: 2012-2013
Department of Statistics
University of Rajshahi
Rajshahi-6205, Bangladesh

Department of Statistics, University of Rajshahi
Rajshahi-6205, Bangladesh
December, 2015

DECLARATION

I do hereby declare that the dissertation entitled "Prediction of Final Stature from a New 3D Two−Stage Growth Model", submitted to the Department of Statistics, University of Rajshahi, Bangladesh, for the Degree of Doctor of Philosophy in Statistics is a unique, completely new and original work of my own. No part of it, at any form, has been submitted to any other university or institute for any degree or diploma on other similar purposes.

Rajshahi
December, 2015

(Md. Abu Shahin)
Ph.D. Fellow
Department of Statistics
University of Rajshahi
Rajshahi-6205, Bangladesh

রাজশাহী বিশ্ববিদ্যালয়
University of Rajshahi

ড. মো: আইয়ুব আলী
প্রফেসর
পরিসংখ্যান বিভাগ
রাজশাহী বিশ্ববিদ্যালয়
রাজশাহী ৬২০৫, বাংলাদেশ
টেলিফোন: +৮৮-০৭২১-৭১১১২২ (অফিস)
মোবাইল: +৮৮-০১৫৫৬৩১০২৯৪
ফ্যাক্স: +৮৮-০৭২১-৭৫০১৯৪
ই-মেইল: ayubali67@yahoo.com

Dr. Md. Ayub Ali
Professor
Department of Statistics
University of Rajshahi
Rajshahi-6205, Bangladesh
Phone: +88-0721-711122 (off)
Cell: +88-01556310294
Fax: +88-0721-750194
E-mail: ayubali67@yahoo.com

CERTIFICATE

This is to certify that the dissertation entitled "Prediction of Final Stature from a New 3D Two–Stage Growth Model" is an original theoretical work done by Mr. Md. Abu Shahin, Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh, for the degree of Doctor of Philosophy in Statistics under my supervision. This study has proposed higher dimensional human growth model, which is a new and advanced work on human growth modeling in the world.

I, also, assure that any part of this dissertation in any form or by any means completely or partially are not submitted for any other degree to any other university or institutions. From this dissertation, one paper has been published in the International Journal of Biometrics and Biostatistics, and another one has been published in the peer-reviewed proceedings of the International Conference on Statistical Data Mining for Bioinformatics, Health, Agriculture and Environment, Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh. Another two papers are under construction from this dissertation.

It is also certified that I have gone through the draft and final version of the dissertation and approved it for submission.

(Dr. Md. Ayub Ali)

Supervisor & Professor
Department of Statistics
University of Rajshahi
Rajshahi-6205, Bangladesh

Dedicated to

My Beloved Parents

and

Teachers who are the Sources of My Knowledge

ACKNOWLEDGMENTS

Firstly, I would like to bow my head and express all praise to Almighty Allah. This is because; He kept me physically and mentally well and fit for completing and making this dissertation to be a success.

I am grateful to the University of Rajshahi, Rajshahi-6205, Bangladesh, for granting me the fellowship.

I express my sincere gratitude to Professor Dr. Md. Nurul Haque Mollah (my previous supervisor), Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh, for his valuable help and suggestions for this study.

I express my inner feelings, deep indebtedness and profound gratitude to my honorable supervisor Professor Dr. Md. Ayub Ali, Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh, for his brilliant guidance, sound advice, perfect direction, cordial cooperation, critical suggestion, unfathomable invaluable help and tolerance which made this work possible. In spite of his various preoccupations, he spent his valuable time by giving instructions to me. I wholeheartedly wish to extend my deepest appreciation to him for reviewing this work. I am really fortunate to be a fellow of him.

I am thankful to Professor Dr. Mohammed Nasser, Ex-Chairman and Professor Dr. Md. Ripter Hossian, Chairman, Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh, for their heartily co-operation, help and inspiration with valuable suggestions during course work and the whole process of completing this research work.

I am indebted to all the teachers of the Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh, for their valuable suggestions, helps and inspiration me at every stage of this work also helped me to shape this thesis paper in the present form.

I am very much grateful to the Academic Committee of the Department of Statistics for giving me the opportunity to prepare the thesis as part of partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

I am grateful to all my friends who have been encouraging me to complete the thesis successfully.

I extend my heartfelt thanks to Internet Search Engine ‘Google’ to find out relevant literatures and solve different kinds of problem or topic to carry out this research work.

Finally, I would like to acknowledge my family members, friends and well-wishers especially to my parents for their tremendous sacrifice, encouragement and continuous inspiration throughout the period of my research, which lead me to all the success in this study.

December, 2015 The Author
Department of Statistics
University of Rajshahi
Bangladesh

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND NOTATION

NOTATION	ABBREVIATION
ARIMA	Autoregressive Integrated Moving Average
BTT	Bock, Thissen and du Toit
CV	Cross Validation
D.S. No.	Data Set Number
FS	Father Stature
HKB	Hoerl, Kennard and Baldwin
ICP	Infancy, Childhood and Puberty
JPA-1	Jolicoeur, Pontier and Abidi Model-1
JPA-2	Jolicoeur, Pontier and Abidi Model-2
JPPS	Jolicoeur, Pontier, Pernin and Sempe
LASSO	Least Absolute Shrinkage and Selection Operator
MS	Mother Stature
MSE	Mean Squared Error
PB	Preece and Baines
RMSE	Root Mean Squared Error
RWT	Roche, Wainer and Thissen
SSC	Shohoji, Sasaki and Cole
S_2	Stature at Age 2
S_3	Stature at Age 3
S_4	Stature at Age 4
S_5	Stature at Age 5
S_6	Stature at Age 6
S_7	Stature at Age 7
S_8	Stature at Age 8
S_9	Stature at Age 9
S_{10}	Stature at Age 10
S_{11}	Stature at Age 11
S_{12}	Stature at Age 12
S_{13}	Stature at Age 13
VIF	Variance Inflection Function

ABSTRACT

The purpose of the present study was to establish higher dimensional growth model and then to predict the final stature from a new 3D two−stage growth model. To check the validity of the models, a secondary longitudinal data on age, weight and stature of Japanese boys and girls were used. The proposed 3D growth model showed more precise than that of the existing 2D growth models. After extracting the final stature and stature at different ages from the well fitted proposed model and eliminating the problem of multicollinearity using 'Forward Stepwise Ridge Regression' and 'Least Absolute Shrinkage and Selection Operator (LASSO)' techniques, this study propose four different equations for predicting final stature with higher precession, validity and stability compared to others. These equations are perfectly applicable in Japanese population. But, the form of the proposed model and its procedures can be applied to others populations.

CHAPTER ONE

INTRODUCTION

"Statisticians work at the interface between Science, Mathematics and Philosophy." - B. Efron and R.J. Tibshirani

CHAPTER 1

INTRODUCTION

1.1 Outline

This chapter includes prelude, importance of the study, aim of the study, concept and terminology, justification of the thesis title and organization of the study.

1.2 Prelude

Anthropometry is the science of measuring the size, weight and dimensions of the human body. Height, weight and skinfold thickness can be used to assess fat stores, adequacy of body weight and risk for chronic disease. Using height and weight, clinicians can compare your current weight with your ideal body weight and usual body weight. A current body height and weight may indicate a compromised nutritional status.

1.3 Importance of the Study

When the stature below the average length 50 cm (Needlman, 2003) of a child is unusual for age, there may be concerned about the individual's final stature. In these circumstances, it is helpful to predict his/her final stature. The prediction may reassure the family or indicate a need for laboratory tests to establish the cause of the unusual growth. Growth failure is a very significant term, used in the medical science to describe a growth rate that is below the appropriate growth velocity (speed) for age. In order to know the growth status of children, it is very important to understand whether a growth failure takes place or not. Prediction of the final stature is important to know the future stature of children and this prediction is based on the mathematical model or equation. Appropriate model is capable of predicting the final stature more accurately. All the existing human growth models consider that stature depends only on age. In fact, stature depends not only on age but also on many other factors, such as body weight, chest circumference, sitting height, genetic factors, maternal illnesses during pregnancy, socio-economic disadvantages during and after pregnancy, social/emotional problems during childhood, poor nutrition and environmental or emotional deprivation and so on. Therefore, a new higher dimensional growth model is essential for better prediction of the final stature.

1.4 Aim of the Study

Accurate predictions of final stature are very important for children who are growing or maturing at usual rates and also for children suffering from diseases, such as hypothyroidism, that can alter their potentials for growth in stature. So, it might have been great concern not only to pediatricians, but also parents having a child with short stature. In order to make better prediction of final stature, the better model is essential. A better model is not always better because of its limitation. It has been changing over time. For this, a good number of researchers used several models for estimating final stature. Some of them are Bayley and Pinneau (1952), Khamis and Guo (1993), Khamis and Roche (1994), Onat (1975), Roche et al. (1975a, b) and Wainer et al. (1978). Most of them predicted the final stature through skeletal age.

Ali and Ohtsuki (2001) predicted the adult stature using the biological parameters as well as different stature variables. Recently, Rahman et al. (2004) proposed more accurate equations for predicting the adult stature using biological parameters and stature variable compare to that from Ali and Ohtsuki (2001).

There are many mathematical ways of parametric and non-parametric modeling the human growth curve for estimating biological parameters and distance curve was attempted. The popular growth models are Gompertz model (Merrell, 1931; Deming 1957), Logistic growth model (Merrell, 1931; Deming, 1957), Jenss model (Jenss and Bayley, 1937), Count model (Count, 1943), Double logistic model (Bock et al., 1973), PB models (Preece and Baines, 1978), ICP model (Karlberg, 1989), Reed model (Berkey and Reed, 1987), SSC model (Shohoji and Sasaki, 1987), JPPS model (Jolicoeur et al., 1988), JPA-1 and JPA-2 models (Jolicoeur et al., 1992), Modified ICP model (Johnson, 1993), BTT model (Bock et al., 1994), Kernel's (Non-Parametric) growth model, Wavelet and Polynomial growth models. It is necessary to select a suitable model to achieve a good prediction. Jolicoeur et al. (1992) declared that, till then, JPA-2 showed the best fit as compared with other structural growth models. Ali and Ohtsuki (2001) and Rahman et al. (2004) were pointed out that the BTT model was better than JPA-2 model. All of these two dimensional growth models considered stature or weight as a function of age only. Necessary of higher dimensional growth model is obvious, but no one have attempted yet due to the difficulties and clumsiness of the estimation process, however, a higher dimensional growth model will definitely improve the precession of the fitting procedures as well as the prediction of the final stature. Thus, the purpose of the study is to:

- propose a higher dimensional growth model,
- pointed out the estimation procedure of the proposed model,
- introduce the procedure of finding velocity, acceleration, local maxima, local minima and saddle point of the proposed model,
- establish the proper validation of the proposed model, and
- finally, establish the estimating equations for predicting final stature for a certain population.

1.5 Concept and Terminology

Prediction: A prediction is a forecast, but not only about the weather. *Pre* means "before" and *diction* means has to do with talking. So, aprediction is a statement about the way things will happen in the future, often but not always based on experience or knowledge. A "prediction" may be contrasted with a "projection", which is explicitly dependent on stated assumptions.

Final Stature: Stature is the distance from the bottom of the feet to the top of the head in a human body, standing erect. Stadiometer is used to measure the stature. The measurement unit usually is centimetres when using the metric system or feet and inches when using the imperial system. Final stature is a stature which is never significantly change during the life. The definition of age at final stature is differing by several researchers (Kato et al., 1998) but most of the researchers suggested that the age at 25 years (Bock et al., 1994), the final stature is occurred.

Three Dimensions (3D): In physics and mathematics, a sequence of *n* numbers can be understood as a location in *n*-dimensional space. When $n = 3$, the set of all such locations is called three-dimensional Euclidean space. It is commonly represented by the symbol $\mathbb{R}^3$. This space is only one example of a great variety of spaces in three dimensions called 3-manifolds. In present study, three variables have been used according to the objective. Hence, the term Three Dimensions (3D) is appeared.

Two Stages: Two stages modeling means the final model is estimated using the estimated value and that value comes from another estimated model.

Growth Model: Growth model is a model which is used to estimate the distance curve as well as the growth parameters. There are many fields such as anthropology, botany, demography, economic, fisheries and zoology have been used as a growth model for modeling their own purpose. In the present study, growth model means human growth model. Fitting to parametric as well as non-parametric growth models are used in order to get the estimate of the biological parameters. Specific methodological approaches are required by the analysis of longitudinal growth data. To establish individual growth patterns and to estimate, so-called, biological parameters of the growth curve are one of the main interests of longitudinal growth studies.

1.6 Justification of the Thesis Title

The title of the thesis is "prediction of final stature from a new 3D two−stage growth model". Let we have a data set that have three variables, namely age, weight and stature of Japanese boys and girls and we want to predict their final stature based on age and weight. Many researchers addressed that after age 25 there will be no increasing in human stature (Bock et al., 1994). Therefore, it is also assumed here that final stature will be attained at age 25. So, weight at age 25 will be predicted first and then in the second step final stature will be predicted. That is, prediction will be performed in three dimensional spaces within two stages –which justifies the title of the dissertation.

1.7 Organization of the Study

This dissertation contains six chapters organized in the following ways:

Chapter 1 is the *Introduction* that includes outline, prelude, importance of the study, aim of the study, concept and terminology, justification of the thesis title, and organization of the study.

Chapter 2 is the *Genesis of the Study* that contains outline, review of literature, research gap and objectives of the study.

Chapter 3 is the *Materials of the Study*. It contains outline, definition of the variables studied, and description and estimation procedure of missing value of data set.

Chapter 4 is the *Methods of the Study*. It contains outline, different methods for the analyzing of data, software and *R* programming code.

Chapter 5 is the *Results and Discussion*. It consists of outline, the research findings and discussion of the compering mean square error of the BTT model and the proposed model, estimate weight for final stature, prediction of final stature, forward stepwise ridge regression model and least absolute shrinkage and selection operator model for selecting appropriate variable with proper diagnostic checks and measures their accuracy.

Chapter 6 is the *Conclusion and Recommendation* which includes concluding remarks consisting of outline, major findings, limitations of the study and scope for the further research.

A bibliography is appended at the end. A sample data set is appended in the appendix-1. The summary of forward stepwise ridge regression model is affixed in the appendix-2.

Title page, declaration, certification, dedication, acknowledgement, table of contents, list of tables, list of figures, list of abbreviations and notation, abstract are appended at the beginning of the dissertation.

CHAPTER TWO

GENESIS OF THE STUDY

"Science is a Wonderful Thing If one Does Not Have to Earn One's Living at It." - Albert Einstein

CHAPTER 2

GENESIS OF THE STUDY

2.1 Outline

This chapter includes genesis of the study, which contains the more important review of literature, research gap and objective of the study.

2.2 Review of Literature

Stature is one of the most important clinical parameters in the identification of an individual, living or dead. It is to be noted that a number of factors such as race, gender and nutrition play a significant role in determining the stature of an individual. When intact bodies are to be examined, stature estimation does not pose any problem. But when dismembered human body parts are the materials to work with, it is a far greater challenge for the forensic pathologists. The history of child growth and development study is not very old. In between 1927-32, several research centers and institutes developed on the multidisciplinary study of child growth of which Fels Research Institute (1929), Institute of Human Development, Berkeley the Child Research Council (Denver), Harvard School of Public Health Growth Study Center (Boston) are important. White House Conference (1933) on Child Health and Protection recommended the need for such studies. The idea behind such studies was partly to protect children from the worst effects of the great depression and partly to acquire further knowledge to determine the effects of the great depression and the possible remedial measures to mitigate these effects. The modeling of human growth is very complicated because of insufficient longitudinal data. Collecting longitudinal data is very costly and time-consuming. Therefore, many researches have stopped due to the unavailability of longitudinal data.

There are various kinds of research on the growth modeling and growth of human body and organs.

Huxley (1932) proposed an allometric equation to find the rate of increment of one dimension of an organ with respect to another of the same or other organ, usually total body size. Huxley pointed out that if two parts (*y* and *x*) grow in accordance with the equation,

$$\frac{dy/dt}{y} = b\frac{dx/dt}{x}$$

After integrating on both sides of the above equation, we got the following form of the equation was as follows:

$\log y = b \log x + \log a$, or $y = ax^b$

This was log linear relationship equivalent to non-linear relationship between *y* and *x*. Here, the dependent variable *y*, which represented a dimension whose increase, was considered relative to that of the independent variable *x*, which might represent a different dimension of the same organ or more commonly a measure of total body size. The logarithmic function could represent a rectilinear plot of the original variables using logarithmic coordinates. Where, '*b*' was the slope of the regression line, which represented the rate of increment of *y* with respect to *x*, and the constant 'log *a*' was the intercept on the *y*-axis, which represents the point of initial growth (Huxley and Tessier, 1936).

Suski and Angeles (1935) was asked by American colleagues, what should be the height and weight of a Japanese child at a given age, but the answer would be that there was not much difference from the height and weight of white American children up to fifteen or sixteen years of age. This statement was based upon the annual measurements of Japanese children in Los Angeles for several years, and the comparison of these figures with American as well as Japanese standards. The children born in America of Japanese parents were found to surpass, between ages of seven and fifteen, Japanese children in Japan, by 7 percent in stature, 20 percent in weight, 9 percent in total leg length, and 7 percent in chest circumference.

Jenss and Bayley (1937) described a four-parameter nonlinear model, namely Jenss model. This model was negatively accelerated exponential and approaches a linear asymptote with positive slope. This model could be written as:

$$y = a_o + a_1 t - \exp(c_o + c_1 t) + \varepsilon$$

where, t was age (years), y was observed stature (cm) or body weight (kg), and ε was random error; a_o, a_1 and c_o were positive parameters, and c_1 was negative.

The $\exp(c_1)$ be the growth or acceleration constant, was independent of scale and measures the ratio of the acceleration of growth at any given age, t to the acceleration at the preceding age, t_{-1} be noted by Jenss and Barley (1937). Thus, to compare the growth of different characteristics within the child, or to study the growth of the same characteristic in different children could be needed by $\exp(c_1)$. The acceleration's magnitude of the constant $\exp(c_1)$ was what largely determines the shape of an individual curve. Since then the model had been used by others (Berkey, 1982; Deming and Washburn, 1963; Manwani and Agarwal, 1973).

Count (1943) proposed a growth model, namely Count model in human stature of Chinese population (children), this model had been applied within the age range three months to seven years. Many other researchers (Tanner et al., 1956; Israelsohn, 1960; Wingerd, 1970 and Mata, 1978) had been applied this model. The linear Count model could be written as:

$$y = a_o + a_1 t + a_2 \ln(t) + \varepsilon$$

where, y was physical measurement (i.e., stature or body weight), t was age (years), ε was random error, and a_o, a_1 and a_2 were the parameters of the model.

The location of zero age was an implicit fourth parameter in the model. Some authors had used conception, or other points that make the interpretation of parameters especially convenient, for age zero.

Bayley and Pinneau (1952) proposed a method to predict the percentage of adult stature achieved. These percentages were provided in tables for chronological age groups of children categorized by whether the Greulich and Pyle (1950) skeletal age differs from the chronological age by more or less than 1 year. This percentage was used with percentage stature to calculate a predicted adult stature.

Deming (1957) and Merrell (1931) discussed in detail the properties of the Gompertz and logistic functions for analyzing human growth process. The Gompertz growth model was used in different discipline, especially, in economic growth and demography. The Gompertz curve could be written as follows:

$$y = p + ke^{-e^{a-bt}}$$

where, y = Dependent variable (i.e., stature), $t =$ Independent variable (i.e., age), p = Lower asymptote (i.e., stature at the start of the adolescent growth cycle), k = Adolescent gain (i.e., stature gain during the adolescent growth cycle), a = Constant of integration (i.e., depending on the position of the origin), b = Rate of constant (i.e., 1/age).

This curve might be considered as the individual's constant rate of maturation through the adolescent growth cycle according to Deming (1957).

And, the mathematical form of the logistic curve was as follows:

$$y = p + \frac{k}{1 + e^{a-bt}}$$

The parameters of logistic growth model had meanings were the same as in the Gompertz function. Marubini et al. (1971) clearly showed the asymmetry of the Gompertz and the symmetry of the logistic curve.

Bock et al. (1973) showed that individual curves for growth in recumbent length from one year to maturity could be represented in good approximation by the sum of two (double) logistic components for the data from the Fels growth study. The first component described growth occurring throughout the pre-pubertal period and continuing in same degree until maturity; the second described the adolescent growth spurt. The double logistic model could be written as:

$$y = \frac{a_1}{1+\exp[-b_1(t-c_1)]} + \frac{f-a_1}{1+\exp[-b_2(t-c_2)]}$$

where, y was the stature (cm), t was the age (years), and a_1, b_1, c_1, b_2 and c_2 were the five parameters. Mature size (f) had to be inserted in the function in this model.

There were six parameters of components, five of which could be estimated by non-linear least squares, and the sixth was the mature stature taken directly from the data. A reliability analysis of the parameter estimates for the Fels samples showed that most of the individual differences in the growth pattern, within sex, could be attributed to three, or at most four, out of the six parameters. Distributions of estimates of these four parameters were presented and discussed in relation to sex differences.

Onat (1975) studied growth and sexual development of 119 normal girls aged 8.5 to 13.4 years were followed from 7 years at 6-month intervals until adult height was reached. The correlation and regression studies showed that the percentage of adult height attained was dependent on age at onset of sexual development as well as the rate of skeletal development. The standard deviations of attained percentage of adult height in relation to age at onset of secondary sexual characteristics were much smaller compared to those based on chronological age and at about the same level as those based on skeletal age. The comparison of the standard errors of methods based on the onset of secondary sexual characters, skeletal age and chronological age showed that the adult height of these girls could be estimated from the height attained at onset of secondary sexual characteristics with an error which was much smaller than that based on height at chronological age and about equal to that based on skeletal age.

Roche et al. (1975a) used the Roche-Wainer-Thissen (RWT) method employs regression equations in which the length and weight of the child, mid-parent stature and Grculich-Pyle (1950) skeletal age, obtained as the median of bone-specific skeletal ages, were used to predict adult stature. The prediction errors with the RWT method were smaller than those with the method of Bayley and Pinneau (Bayley and Pinneau, 1952) and that was conformed when applied to data from three longitudinal growth studies.

Preece and Baines (1978) proposed a new family of mathematical functions to fit longitudinal growth data and developed a procedure for fitting individual serial record of stature from age two to adulthood and also described the properties of biological parameters of their proposed growth model. The model was as follows:

$$H(t,\theta) = h_1 - \frac{2(h_1 - h_2)}{\exp[s_o(t-\delta)] + \exp[s_1(t-\delta)]}$$

where, $H(t,\theta)$ was the stature (cm) at age t, θ was a growth parameter vector $(h_1, h_2, s_o, s_1, \delta)$, h_1 was the equation parameter, which was the estimated adult stature. The parameters s_o and s_1 were rate constants, and h_2 and δ were related to the stature and age at take-off of the adolescent growth spurt.

Several authors (Billiwicz and McGregor, 1982; Bogin et al., 1990; Bogin et al., 1992; Brown and Townsend, 1982; Byard et al., 1993; Cameron et al., 1982; Hauspie, 1980; Hauspie et al., 1980a; Hauspie et al., 1980b; Jolicoeur et al., 1988; Jolicoeur et al., 1992; Ledford and Cole, 1998; Mirwald et al., 1981; Qin et al., 1996; Tanner et al., 1982; Zemel and Johnston, 1994) were also used the above model.

Wainer et al. (1978) discussed the RWT method for predicting adult stature from childhood variables used the current recumbent length and weight of the child, the stature of each parent, and the skeletal age of the child as predictor variables. There was only a small increase in the errors of prediction if population mean values were substituted in the prediction equations when the father's stature, the skeletal age of the child, or both these variables were unknown. This modified method was more generally applicable than the original RWT method.

Onat (1983) used multivariate regressions for estimating adult height that were presented based on height, skeletal age (SA), chronological age and mid-parental stature (MPS) of Turkish prepubescent or early adolescent girls. Discriminating these regressions with regard to information on the presence or absence of secondary sex characters, as well as menarche improves the estimations, especially in those who developed either early or late in respect to secondary sexual development and in postmenarcheal girls. The 3-variable regressions, neglecting sexual maturity, resulted in over estimation of adult height in early maturing girls. These were corrected by regressions in which the states of secondary sex characters were used as dummy variables in addition to height, SA and MPS.

Berkey and Reed (1987) proposed a new models, namely Reed model which was appropriate for early childhood growth in length and possibly also for weight and head circumference of 229 Baston children. These models were actually the extension of the Count model (Count, 1943). The first-order Reed model could be written as follows:

$$y = A + Bt + C\ln(t) + \frac{D}{t}$$

The first-order Reed model had four parameters and it was more flexible than the Count model since it allows an inflexion point. The second-order Reed model could be written as:

$$y = A + Bt + C\ln(t) + \frac{D}{t} + \frac{E}{t^2}$$

In the second-order Reed model, the fifth parameter allowed a second inflexion point. The first-order version was shown to perform well on height between 3 months and 6 years but few children needed the second-order version (Berkey and Reed, 1987). An extra benefit of this model was that a wider variety of both normal and abnormal growth patterns could be accommodated by the curve.

Shohoji and Sasaki (1987) described a growth model, which had six parameters. It could be written as the following form:

$$y(t) = AW(t) + f(t)[1 - W(t)] + \varepsilon$$

where, t was postnatal age, $y(t)$ was stature at age t, A was adult stature, $W(t)$ was a weighting function given by $W(t) = \exp[-\exp\{B(G-t)\}]$, $f(t)$ was a function of stature in infancy given by $f(t) = C + Dt + E\log t$ and ε was an error.

The weighted average of adult stature A was the stature at age t and stature predicted from an infancy model $f(t)$. The Gompertz function was the weight $W(t)$ takes the value zero at $t=0$, then switches from 0 to 1 at G, with parameter B controlling the suddenness of the switch. The function $f(t)$ was the Count model for infant stature and body weight. However, the Jenss-Bayley function $f(t)=C+Dt-\exp(E-Ft)$ was another infant stature model with one extra parameter, combining an exponential and a linear component, which performed appreciably better (Berkey, 1982), suggested modifying the Shohoji-Sasaki model to use the Jenss-Bayley rather than the Count model as its childhood component. Another seven-parameter model (KS7) was described by Kanefuji and Shohoji (1990) extending that of Shohoji and Sasaki (1987), replacing the Count model by $f(t)=C+Dt+\log(E+Ft)$. This combined of an exponential infancy, linear childhood and logistic puberty component produced the SSC (Shohoji-Sasaki modified by Cole) model which was similar to Karlberg's ICP model (Karlberg, 1989; Ledford and Cole, 1998).

Jolicoeur et al. (1988) proposed a seven-parameter asymptotic growth curve, namely JPPS model that had been applied to longitudinal data on the height of 13 boys and 14 girls from 1 month to 19 years of age. This new curve was expressed with respect to total age, passes through the origin, and fits infants as satisfactorily as older children. The form of this model could be written was as follows:

$$y(t)=A\left\{1-\frac{1}{1+\left(\frac{t'}{D_1}\right)^{C_1}+\left(\frac{t'}{D_2}\right)^{C_2}+\left(\frac{t'}{D_3}\right)^{C_3}}\right\}+\varepsilon$$

where, t' was post-conceptual age; $y(t)$ was stature at age t'; A was adult stature; D_1, D_2 and D_3 were positive age scale factors; C_1, C_2 and C_3 were positive dimensionless exponents; and ε was the error. Note that t' was age post-conception, i.e., $t' = t + 0.75$ assuming a constant gestation of 9 months.

The residual sums of squares with this new curve were 7.5 times lower on the average than with the currently-used five-parameter curve of Preece and Baines (1978) and 2.4 times lower than with the six-parameter curve of Shohoji and Sasaki (1987).

Karlberg (1989) developed the ICP model and this model divided growth into three distinct phases of functional form, such as, Infancy, Childhood, and Puberty. These three distinct phases of functional form were described below:

An Infancy component assumed to start during fetal life with a rapidly decelerating course ceasing at 3–4 years of age and also it was explained by an exponential function:

$$y = a_1 + b_1\{1 - \exp(-c_1 t)\}$$

A Childhood component started during the first year of life having a slowly decelerating course and continuing until end of growth. A second degree polynomial function explained this component and this polynomial function could be written as:

$$y = a_c + b_c t + c_c t^2$$

A Puberty component representing the additional growth induced by puberty and accelerating up to age at peak velocity (age = t_V), then decelerating until the end of the growth (age = t_E). A logistic function represented this component and that function was:

$$y = \frac{a_p}{1 + \exp\{-b_p(t - t_V)\}}$$

In all the above three functions y was stature for the relevant component at time t in years from birth, and t_E was the middle of the first one year interval after age at peak velocity where the overall gain becomes less than that in the Childhood component.

These components of the human growth curve from birth to adulthood strongly reflected the different hormonal phases of the growth process. This model provided an improved instrument for detecting and understanding growth failure.

Lindgren and Hauspie (1989) conducted a longitudinal study of physical growth of Swedish School Children, born in 1955 and aged 10 to18 years. They showed the secular changes in height, weight and weight–for–height as expressed in BMI. They compared average heights and weights over 10-15 years for the samples of Swedish School Children born in 1955 and 1967. They showed that both boys and girls had been gaining more weight than height, especially around the ages at which peak velocity generally occurs. Since the increasing height of children born in 1967 gradually diminished after age at peak height velocity, it seems that the height difference during puberty mainly reflects an earlier maturation of these child's, compared to the child's born in 1955.

Tsuzaki et al. (1990) studied the difference in head circumference between Japanese and Caucasian Children. The subjects consisted of a total of 42392 Japanese children between zero and four years of age surveyed from 1940 to 1980, and those data were compared with those of American and British children. They found that, there was a significant ethnic difference in head circumference, as large as one channel of usual percentiles, between Japanese and Caucasian children. The results indicated that smaller head circumference in Japanese children primarily reflects smaller stature of the Japanese.

Feldesman (1992) examined the relationship between femur length and stature in children between the ages of 8 and 18 years. They showed that the femur stature ratios of children between the ages of 8 and 11 differ significantly from their older counterparts. Between the ages of 12 and 18, there were no significant differences due to age in the femur/stature ratio; however, there were significant differences in this age group attributable to gender. They also showed that the worldwide average adult femur/stature ratio does not adequately describe children in this age range. Their study strongly documents the adolescent growth spurt in the femur/stature ratios of both males and females at the precise time one would expect to see the spurt occur (10-12 in females; 12-14 in males). This growth follows a nearly identical trajectory in both genders, with relative femur growth dominating before the peak years of the growth spurt, and relative stature growth dominating afterward. This accounted for the ratio’s rise to maximum values just before peak growth, and it’s declined toward the adult ratio thereafter. These findings required us to use separate adolescent femur stature ratios of 27.16 (females) and 27.44 (males) to estimate the stature of children between the ages of 12 and 18. Preliminary testing showed these ratios to be more accurate in estimating stature than the properly selected Trotter and Gleser adult regression equation. Used of the adolescent male ratio with the Homo erectus juvenile WT 15000 resulted in a lower stature estimate (157.4 cm) than previously reported. It was suggested that continued testing of the ratio occur, but that the values here in derived might be useful in routine forensic cases involving children in this age range, and with sub-adult paleontological specimens.

Jolicoeur et al. (1992) discussed the eight different asymptotic models for the comparing with respect to their goodness of fit for the description of the longitudinal growth of stature in 27 healthy children from the French Auxological Survey. Some growth models were based on total age (measured from the time of fertilization) and some were based on postnatal age (age after birth). Their results showed that some of these models were the most accurate, but they would not be suitable for prenatal data or extrapolations. Finally, they proposed two models which were extensions to the model of Jolicoeur et al. (1988) where the age offset was estimated from the data rather than being constrained at 0.75, to improve the fit in infancy. The extended models were as follows:

$$y(t) = A\exp\left\{-\frac{1}{c_1 \log_e\left(1+\frac{t}{D_1}\right)+\left(\frac{t}{D_2}\right)^{C_2}+\left(\frac{t}{D_3}\right)^{C_3}}\right\}+\varepsilon$$

$$y(t) = A\left\{1-\frac{1}{1+\left(\frac{t+E}{D_1}\right)^{C_1}+\left(\frac{t+E}{D_2}\right)^{C_2}+\left(\frac{t+E}{D_3}\right)^{C_3}}\right\}+\varepsilon$$

JPA-1 was the nickname of the model in first equation and JPA-2 was the nickname of the model in second equation. The models JPA-1 retained the theoretically desirable quality of passing through the origin with respect to total age while, JPA-2 fitted human stature data better than all other asymptotic models proposed till 1991 (Jolicoeur et al., 1992).

Ashizawa et al. (1993) estimated the diversity of adolescent growth, spline-smoothed individual velocity curves of stature, body weight and chest circumference of 44 girls in Tokyo, of which menarche was recorded correctly. Additionally, 25 variables of ages at peak velocity, intensities, sizes and weight at the peak and at menarche, and terminal height were obtained.

Johnson (1993) proposed the modified ICP model which was used to convert the non-stationary time series of growth observations into a stationary time series for the Fourier analysis. This model described the combined form of two phases such as Childhood and Puberty. The modified form was as follows:

$$y_i = a_C + b_C t_i + \frac{a_p}{1 + \exp\left\{-b_p\left(t_i - t_V\right)\right\}}$$

Khamis (1993) studied Roche-Wainer-Thissen (RWT) prediction model for the predicting the adult stature of a child based on age, current stature, current weight, current skeletal age and the average stature of the parents, and found out an improvement of this prediction equation. They investigated the seven variations of the current version of the RWT prediction model and compared in terms of the accuracy and reliability of prediction, culminating in a recommendation for the prediction of adult stature in Caucasian Americans. Their proposed method, called multivariate cubic spline smoothing, used cubic splines in the smoothing part of the RWT prediction model, resulting in a simpler (i.e., fewer steps) method with smaller maximum deviations between predicted and actual adult statures than the current multivariate semi-metric smoothing method.

Takai (1993) studied 6300 Japanese children from Ogi Growth Study 1979-1988 for the describing the velocity of the Tanner-Whitehouse 2 skeletal maturity. The cubic *B*-spline function fitted the velocities for the Carpal, RUS and 20-bone scores on the smoothed velocity curves. The maturity velocity curves showed single peak around the adolescent period exclusive of a bimodal curve for girls' RUS velocity. Its first peak appeared at 10.9 years and the second, 13.9 years. Just after the first peak their height reached the maximal velocity (11.0 years). The RUS velocity curve for the boys showed the peak maturity velocity at 15.6 years. Their height attained the peak velocity (at 12.9 years) before the RUS maturity did. The study also showed that the skeletal maturation affected the height growth during the duration of height spurt for the boys, but only during the accelerating period for the girls.

Bock et al. (1994) described the triphasic generalized logistic model by summing up three phases of growth; early, middle and adolescent. This model was popularly known as BTT (<u>B</u>ock-<u>T</u>hissen-du <u>T</u>oit) model. This triphasic generalized logistic model could be written as:

$$y = \frac{a_1}{\left[1+\exp\left(-b_1 t - c_1\right)\right]^{d_1}} + \frac{a_2}{\left[1+\exp\left(-b_2 t - c_2\right)\right]^{d_2}} + \frac{a_3}{\left[1+\exp\left(-b_3 t - c_3\right)\right]^{d_3}}$$

where, the set of parameters (a_1, b_1, c_1), (a_2, b_2, c_2) and (a_3, b_3, c_3) referred to the parameters of early, middle and adolescent phases of growth, respectively. And, d_1, d_2 and d_3 were the fixed shape constant to the respective phases.

Khamis and Roche (1994) used longitudinal data (every 6 months) from participants residents of Southwest Ohio, 223 white males and 210 white females, at the time of their enrollment into the Fels Longitudinal Study. They applied a modification of the Roche-Wainer-Thissen stature prediction model to obtain reliable and accurate predictions of adult stature in white American children who were free of disease without using skeletal age. They concluded that adult stature predictions were needed commonly but the current methods were difficult to apply because they require a skeletal age assessed by a modem method. The Khamis-Roche method predicted adult stature in the absence of skeletal age with only a slight deterioration in accuracy and reliability. The applicability of the Khamis-Roche method was limited to white American children without pathologic conditions that alter the potential for growth in stature, but it should be useful for white children who were unusual in stature or in levels of maturity for age.

Broeck et al. (1995) used 153 patients with Turner syndrome, padicipating in five European trials, were included to study final height after long-term growth hormone (GH) treatment in girls with Turner syndrome (TS). They found that at the last measurement, mean (SD) height was 150.7 (4.9) cm in group 1 and 148.5 (5.1) cm in group 2. The differences between final height (FH) and projected final height based on extrapolation of the initial height-standard deviation score on Turner syndrome reference values, were 2.9 (3.8) and 3.0 (3.3) cm, respectively. The mean gain over the Bayley-Pinneau prediction of FH was 3.3 (3.9) cm in both groups. They were found that no significant differences between countries. The range of gains over projected height (-4.7 to 12.1 cm) was large, and 25% of gains were 5 cm or more. Gain over initial projection was strongly related to initial growth delay and to growth response during the first 2 years of treatment. A logistic regression model was presented that predicts gain of more than 5 cm with a positive predictive value of 62% and a negative predictive value of 84%. And, they concluded that long-term GH treatment in girls with TS, starting treatment at a relatively advanced age (>10 years) resulted in a modest mean gain in FH of 3 cm, with wide inter individual variation.

Rosique and Rebato (1995) studied on regional differences in the growth of Spanish Children by fitting the Preece-Baines Model 1 to cross sectional stature data. They compared children from seven different studies using the function parameters and derived biological variables. Regional differences in growth were interpreted as a result of a geographic variation among Spanish provinces in demographic, public health and nutritional conditions. Adult stature and the pattern of growth differed between urban and rural populations from the interior lands. Males from urban Extremadura, Barcelona and the Basque country showed that the tallest adult statures. Adult statures of males from Segovia, Extremadura emigrants and Cuenca were not only the lowest but the growth pattern showed delay in estimated ages at take-off and PHV compared to the other populations. Estimated age at PHV was later for all male samples compared to Vizcaya, except for the sample of Barcelona-I. Females from Barcelona-II, Segovia and the Basque Country showed the tallest adult statures. All of the female samples, except that of urban Extremadura, had an earlier estimated age at PHV compared to the sample from Vizcaya.

Tibshirani (1996) proposed a Least Absolute Shrinkage and Selection Operator (LASSO) method for estimation in linear models which minimizes the residual sum of squares subject to the sum of the absolute value of the coefficients being less than a constant. Because of the nature of this constraint it tends to produce some coefficients that were exactly zero and hence gives interpretable models. Their simulation studied suggests that the LASSO enjoys some of the favourable properties of both subset selection and ridge regression. It produced interpretable models like subset selection and exhibits the stability of ridge regression. The LASSO idea was quite general and could be applied in a variety of statistical models: extensions to generalized regression models and tree-based models were briefly described.

Chumlea et al. (1998) collected anthropometric data for stature, knee height, and sitting height from a gender and racial/ethnic-stratified sample of 4750 persons from the US population (1369 non-Hispanic white men, 1472 non-Hispanic white women, 474 non-Hispanic black men, 481 non-Hispanic black women, 497 Mexican-American men, 457 Mexican-American women) aged 60 years or older participated to develop new, nationally representative equations to predict stature for racial/ethnic groups of the elderly population in the United States. They used sampling weights to adjust the individual data to account for unequal probabilities of selection, nonresponse, and coverage errors so that all individual data used in these analyses represented national probability estimates. Regression analysis was performed to predict stature in each gender and ethnic group, and the results were cross-validated. Stature prediction models using knee height and age and sitting height and age were evaluated for each gender and racial/ethnic group. The equations with knee height and age were selected on the basis of root mean square error and pure errors in cross-validation and on the accuracy and validity of measures of knee height over sitting height. New stature prediction equations using knee height and age were presented for non-Hispanic white, non-Hispanic black and Mexican-American elderly persons from current nationally representative data. These equations should be applied when a measure of stature couldn't be obtained.

Henneberg and Louw (1998) discussed on the patterns of physical growth (height, weight, length of body segments, circumference and widths) and functions (grip strength, reflexes and pulse rates) "Cape Coloured" School children of the data on selected urban and rural groups with maximum contrasting socio-economic status (SES). They showed that the heights and weights of pre-pubertal urban children match American reference data, but post pubertally they decline somewhat, whereas these measurements of the rural children consistently lie~1 standard deviation below the urban group. Skin folds thickness of urban children match or exceed the American reference, implying that their nutritional needs were being met well. Functional indicators of rural children were much poorer than those of urban children.

Kato et al. (1998) examined the three definitions of final height applied to the data were: (1) Final stature at 18 years of age; (2) Stature after a year with an annual increment less than 0.5 cm; and (3) The highest measurement, and their validity when practically applied to two different longitudinal data sets they were T-data (31 boys and 35 girls born between 1967 and 1978) and H-data (113 girls born between 1956 and 1966). Their results suggested that the greatest height of an individual measurement was the most effective definition of 'final stature' for practical use. This definition could be applied to various types of data, whether measurements were obtained from individuals during school periods, or whether measurements were obtained from individuals until the cessation of growth.

Ali and Ohtsuki (2000) estimated the maximum increment age (MIA) in height and weight of Japanese boys and girls during the birth years 1893-1990 using the published data of the ministry of education, science, sports and culture in Japan. They found that the estimated MIA showed an overall declining trend, except in birth year cohorts in 1934-1951.

Cole (2000) compared the two midparent height calculations, and to see if they explained the imputation procedure used by Galton to adjust for the difference in adult height between daughters and sons. Galton multiplied daughters' heights by 1.08 before averaging them with sons' heights. Using data from 17 national height references they shown that this procedure was equivalent to averaging the height standard deviation scores (SDSs) of the sons and daughters. It demonstrates that midparent height SDS obtained by averaging the height SDSs of the two parents was a valid alternative to conventional midparent height.

Ali and Ohtsuki (2001) analyzed longitudinal growth in stature for 509 males and 311 females of Japanese from childhood to adulthood. They were extracted the growth parameters from estimated distance and velocity curve for each individual using triphasic generalized logistic (BTT) growth model that was done AUXAL software program. The forward stepwise regression model was used to predict adult stature based on biological parameters and found that the adult stature depend on SPHV and STO for both boys and girls who had with and without mid growth spurt.

For boys (whole sample)

$$PAS = 1.534135\ SPHV - 0.483147\ STO$$

For boys (with mid-growth spurt)

PAS = 1.539818 SPHV - 0.498201 STO

For boys (without mid-growth spurt)

PAS = 1.558884 SPHV - 0.503437 STO

For girls (whole sample)

PAS = 1.628559 SPHV - 0.580845 STO

For girls (with mid-growth spurt)

PAS = 1.488222 SPHV - 0.431707 STO

For girls (without mid-growth spurt)

PAS = 1.592562 SPHV - 0.534061 STO

Leigh (2001) discussed the human pattern of growth and development appears to differ markedly from patterns of ontogeny in other primate species. Humans present complex and sinuous growth curves for both body mass and stature. Many human proportions changed dramatically during ontogeny, as we reach sizes that were among the largest of living primates. Perhaps most obviously, humans grow for a long time, with the interval between birth and maturation exceeding that of all other primate species. These ontogenetic traits were as distinctive as other key derived human traits, such as a large brain and language. Ontogenetic adaptations were also linked to human social organization, particularly by necessitating high levels of parental investment during the first several years of life.

Fukami et al. (2003) studied of longitudinal auxological in a 14 year 9 month old Japanese girl with Lćri–Weill dyschondrosteosis accompanied by mesomelic short stature, who had a submicroscopic pseudoautosomal deletion involving SHOX, and pubertal development of an almost average tempo to report on auxological data in the combination of SHOX (short stature homeobox containing gene) haploinsufficiency and normal ovarian function. The standard deviation scores (SDSs) for height, leg length (LL), and arm span remained below the normal range from childhood and worsened during puberty, whereas those for sitting height (SH) remained within the normal range and stayed almost constant throughout the observation period. Consequently, the SDSs for SH/LL ratio remained above the normal range from childhood and deteriorated during puberty. The decreased pubertal height gain was caused by a diminished pubertal height spurt and abrupt growth cessation shortly after menarche. The SDSs for hand length and palm length remained within the normal range but decreased during puberty, and those for head circumference remained within the normal range and stayed almost constant throughout the observation period. Their results suggested that, in individuals with SHOX haploinsufficiency and normal ovarian function, auxological abnormalities related to mesomelia were evident from childhood and worsen further during puberty because of the skeletal maturing effects of ovarian estrogens.

Shahar and Pooy (2003) developed an equations using several anthropometric measurements of a cross sectional study such as body weight, height, arm span, half arm span, demi span and knee height of 100 adults (aged 30 to 49 years) and 100 elderly subjects (aged 60 to 86 years) from three major ethnic groups of Malays (52%), Chinese (38.5%) and Indians (9.5%) participated for estimating stature in Malaysian elderly. The %CV of anthropometric measurements in adults and elderly subjects ranged between 5 to 6%, with standing height having the lowest %CV. When the equations derived from adults were applied to elderly subjects, it was found that percentage difference between actual height and the estimated value ranged from 1.0 to 3.3%. However, the percentage difference between estimated heights from the equations developed in this study compared to those derived from the equations of other populations ranged between 0.2 to 8.7%. They concluded that the standing height was an ideal technique for estimating the stature of individuals.

Ali et al. (2004a) analyzed longitudinal growth of stature for 509 boys and 311 girls from early childhood to adulthood to predict the average adult stature. They were extracted biological variables from the fitted triphasic generalized logistic model (BTT model) and this model was estimated by AUXAL software using the estimated population mean values and covariance matrix values for the Japanese population. They found that significant inter-correlations among the biological variables. Japanese boys and girls were characterized by earlier age at peak height velocity and shorter stature with medium peak height velocity, that were comparing with other populations; the parameters in the BTT model decomposed that, on average, 47.8%, 38.7%, and 13.5% of the adult stature were completed respective during the early, middle and adolescent growth phases, for the Japanese boys. For the girls, these percentages were 44.0%, 42.9%, and 13.1%, respectively. Also, they found that the average predicted adult stature of Japanese boys was 172.59 cm and that of girls was 159.68 cm for Japanese population.

Ali et al. (2004b) used the stepwise regression approach to predict the final stature of Japanese children from the distance curve using the sample of 509 boys and 311 girls. After removing the outliers and influential data points, regression equations were highly cross validated, and they proposed prediction equations for the final stature of Japanese boys and girls, separately. Finally, they proposed equations of predicting final stature for the Japanese were as follows:

For boys (average):

$PFS = 1.27066\ S_9 + 0.63875\ S_3 – 0.377094\ S_{12}$

For boys (who had mid-growth spurt)

$PFS = 1.281532\ S_9 + 0.493674\ S_3 -0.296906\ S_{12}$

For boys (who don't had mid-growth spurt)

$PFS = 1.080969\ S_9 + 0.730981\ S_2 – 0.304324\ S_5$

For girls (average)

$PFS = 3.20074\ S_{13} – 3.29566\ S_{12} + 1.11848\ S_{11}$

For girls (who had mid-growth spurt)

$PFS = 3.09039\ S_{13} – 3.23795\ S_{12} + 1.16692\ S_{11}$

and For girls (who don't had mid-growth spurt)

$PFS = 3.18652\ S_{13} – 3.03767\ S_{12} + 0.86965\ S_{11}$

Rahman et al. (2004) used 483 males and 262 females' longitudinal data to fit double phasic growth (JPA-2) and triphasic generalized logistic (BTT) models through the software AUXAL 2.01 for characterized individual growth of stature and find out more efficient prediction equations. They growth parameters extracted from the estimated distance and velocity curves for each individual. Six prediction equations of adult stature on growth parameters (an improvement of Ali-Ohtsuki equations) had been established for Japanese boys and girls.

For boys (whole sample individuals)

$PAS = 1.046828\ S_9 + 0.397943\ S_3$

For boys (who had the mid-growth spurt)

$PAS = 0.881007\ S_{10} + 0.543971\ S_4$

For boys (who do not had a mid-growth spurt)

$PAS = 1.009736\ S_9 + 0.412471\ S_4$

For girls (whole sample individuals)

$PAS = 2.87668\ S_{13}\text{-}1.89627\ S_{12}$

For girls (who had the mid-growth spurt)

$PAS = 2.54599\ S_{13}\text{-}1.55950\ S_4$

For girls (who do not had a mid-growth spurt)

$PAS = 2.63537\ S_{13}\text{-}1.65255\ S_{12}$

Ashizawa et al. (2005) studied the longitudinal growth of Japanese subjects performed by applying the Preece–Baines model 1 (PB1) function. Ninety-three sets of longitudinally-followed height data from a series of girls in Tokyo were analyzed by fitting the PB1 model. They first compared the PB1 results with those previously obtained from the cubic spline function. They were then examined correlation among biological variables within this Japanese group, and then they compared the PB1-derived biological variables among populations. They found the following results. In comparison with previous results obtained by the cubic spline function for the same subjects, the PB1-derived velocity curve was found to be more emphasized. Ages at take-off and at peak were 0.2 years younger and older, respectively, and height was 1 cm less at take-off and 1.3 cm greater at peak. Many variables were significantly correlated within the Tokyo girls. Though, the number of variable pairs that showed a significant correlation was considerably smaller in the among-population comparisons than in the within-group analysis of the Tokyo girls.

Jones et al. (2005) reviewed some of the possible biological maturity indicators that the pediatric exercise scientist could use. As a result, they recommend that any of the methods discussed could be used for gender-specific comparisons. Gender-comparison studied should either use skeletal age or some form of somatic index.

Sunil et al. (2005) analyzed the regression analysis using the variables as height and hand length of 150 healthy individuals (75 males and 75 females) in various colleges of Delhi. Bilateral asymmetry in hand measurements were statistically insignificant. Regression equations were derived for right and left hand separately by which living stature might be fairly accurately estimated when a fragmentary or mutilated portion of upper extremity was recovered. Using the regression formula derived in this study, stature could be estimated within the error of +4.0 to 4.6 cm from hand length.

Csukás et al. (2006) analyzed six longitudinally followed somatometric traits such as height, sitting height, iliospinal height (B–ic), upper limb length (a–da), biacromial diameter (a–a), and biiliocristal diameter (ic–ic) of Japanese boys of Ogi Growth Study for the mathematical growth modeling of Preece and Baines model. Biological variables derived from the estimated parameters were studied with emphasis on duration and velocity characteristics of the adolescent spurt. Ages for measurements at peak velocities tend to be younger than previously reported non-Japanese ones. Spurt duration in limb measurements was significantly the shortest. Earlier age at minimal velocity (AMV) and later age at peak velocity (APV), thus the longest spurt duration, were the characteristic for transverse measurements (a–a, ic–ic). B–ic and a–da had the largest, while a–a and ic–ic had the smallest relative velocity at AMV. Another result for the transverse measurements was that the magnitudes of differences between relative minimal and peak velocities (RMV, RPV) were the largest. They suggested that a high level of RMV results from early maturation of bones, thus leading to the shortest spurt duration in limb dimensions, while a low level of RMV results from late maturation of the bones, consequently leading to the longest spurt duration in transverse measurements. This tendency of reverse relation was present in the rest of the measurements as well. Transformation of velocity variables (minimal velocity (MV), peak velocity (PV)) to relative ones, proved to be useful in observing the relation of spurts in measurements.

Maijanen and Niskanen (2006) compared the stature estimation methods on osteological material from medieval Westerhus, Sweden. They used a recently revised anatomical technique (Raxter et al., 2006) to estimate the living stature (XSTAT) of the individuals that compared with other anatomical methods and various regression equations on long bone lengths to examine their applicability to this skeletal sample and the accuracy of their estimates. They found considerable differences in estimates between techniques, especially in mean statures of tall and short stature classes based on long bone lengths. Thus they emphasized the importance of choosing the most appropriate estimation methods.

Sarajlić et al. (2006) developed appropriate stature estimation formulae from the length of the femur, tibia and fibula from 50 male cadavers, of individuals who died between the ages of 23 to 54 years for use in the Bosnia and Herzegovina to help in identifications of the victims. The cadaver length was measured and the length of the long bones was obtained from *X*-ray photographs. The length of the cadavers of the individuals who died after age of 45 years was corrected according to Giles' table. This study established that using Trotter and Gleser's formulae under estimate stature of tall people in the current population of Bosnia and Herzegovina. Smallest standard error of estimate was observed in the formula that uses the sum of the length of femur and fibula. There were no statistically significant differences between the length of the bones from the left and right sides of the body. Therefore, formulae developed from the average length of bone pairs were recommended for use.

Mounir et al. (2007) used anthropometric assessment including weight, height, mid upper arm circumference (MUAC), waist circumference, hip circumference and triceps skin-fold thickness of sample 1606 girls was conducted in primary and preparatory schools in Alexandria to assess the mean age of menarche and the main nutritional factors affecting it. BMI and body fat percentage were calculated. A 24 hours diet recall method was used to assess the dietary intake. They found the following results. The mean age of menarche was 11.98±0.96 years. The mean MUAC, triceps skin-fold thickness, waist circumference and hip circumference were significantly higher among menstruating girls as compared to non-menstruating. ($p< 0.01$). Only 7.5% of the females less than the 5th percentile of BMI (thinness) were menstruating, while the corresponding figure for those at or more than 85th percentile (overweight) was 65.6% and this was statistically significant (χ_2^2 =102.8, $P < 0.001$). Girls who attained menstruation demonstrated a higher significant mean percent of body fat (43.40±10.0) as compared to non-menstruating ones (35.41±7.87), ($t = 17.09$, $P < 0.001$). The oldest age at menarche was noted when the protein, iron and caloric intake was less than 80% of the RDAs. However, after adjustment of other variables direct relation was detected between age of girls and their age of menarche and those in private school had earlier age of menarche than those in governmental one. The nutritional status of the adolescents had a significant association with the onset of menstruation and the age at menarche.

Stovitz et al. (2008) used 2802 subjects from the Child and Adolescent Trial for Cardiovascular Health (CATCH) to examine the interaction of childhood height and childhood BMI in the prediction of young adult BMI. The associations and interactions between height (cm) and BMI (kg/m2) were assessed using mixed linear regression models with adult BMI as the dependent variable and they found that a significant interaction between childhood height and childhood BMI in the prediction of adult BMI ($P < 0.0001$). Stratification by Centers for Disease Control and Prevention (CDC) reference quintiles revealed that a positive association between childhood height and adult BMI existed only for those subjects in the top quintile of childhood BMI, within whom predicted adult BMI ranged from 27.5 (95% confidence interval = 26.4–28.6) for those in the shortest height quintile to 30.2 (95% confidence interval = 29.7–30.6) for those in the highest height quintile. Among children with high BMI levels, those who were taller, as compared to those who were shorter, had significantly higher young adult BMI levels. This pattern seems primarily due to the positive association of childhood height and childhood BMI. Clinicians should recognized the risk of excess body weight in young adulthood for all children who had a high BMI, and pay special attention to those who were tall, because their childhood height would not protect them from subsequent weight gain and elevated BMI.

Bogin and Silva (2010) discussed the decomposing stature into its major components was proving to be a useful strategy to assess the antecedents of disease, morbidity and death in adulthood. Human leg length, sitting height and their proportions were associated with epidemiological risk for overweight, coronary heart disease, diabetes, liver dysfunction and certain cancers. Human beings followed a cephalo-caudal gradient of growth, the pattern of growth common to all mammals. A special feature of the human pattern was that between birth and puberty the legs grow relatively faster than other post-cranial body segments. For groups of children and youth, short stature due to relatively short legs was generally a marker of an adverse environment. The development of human body proportions was the product of environmental × genomic interactions, although few if any specific genes were known. The HOXd and the short stature homeobox-containing gene (SHOX) were genomic regions that might be relevant to human body proportions. However, research with non-pathological populations indicates that the environment was a more powerful force influencing leg length and body proportions than genes. Leg length and proportion were important in the perception of human beauty, which was often considered a sign of health and fertility.

Ilayperuma et al. (2010) used 258 subjects with an age span of 20-23 years to investigate the relationship and to propose a gender and age specific linear regression models between the ulna length and height of an individual. Their findings indicated that the significant differences of the ulna length between the genders. They found a significant positive correlation between height and ulna length. Regression equations for stature estimation were formulated using the ulna lengths for both males and females. The ulna length provided an accurate and reliable means in estimating the height of an individual. They proposed regression formulae that would be useful for clinicians, anatomists, archeologists, anthropologists and forensic scientists when such evidence provides the investigator the only opportunity to gauge that aspect of an individual’s physical description.

Hui et al. (2011) used the data on height, weight and chest circumference obtained from two serial national cross-sectional surveys for children aged 0 to 7 years in China, to describe the secular trends on physical growth of children during the year of 1985-2005 and to analyze the urban-suburban-rural difference and its change. They showed that the average weight and height for both boys and girls from urban, suburban and rural areas had significantly increased in most age groups during the past 20 years; the average chest circumference increased slightly, ranging from 0.0 to 2.0 cm. From 1985 to 2005, the urban-suburban difference in height had become smaller, and that in weight showed similar trend for children under 3 years old but became larger after 3 years old; the suburban-rural difference both in height and weight became larger after 6 months old. The increment per decade in height was the greatest in the suburban group while the greatest increment in weight was the urban group. They concluded that the positive secular trends were observed among urban, suburban and rural areas in Chinese children under 7 years old during the 1980s and the 2000s, reflecting a rapid socio-economic development in China.

Johnson et al. (2011) investigated the secular trends in weight and length growth from birth to 3 years of age in 620 infants (318 boys and 302 girls) born from 1930 to 2008, and to assess whether these trends were associated with concurrent trends in pace of infant skeletal maturation and maternal body mass index. Their results showed that the most pronounced differences in growth occurred in the first year of life. Infants born after 1970 were approximately 450 g heavier and 1.4 cm longer at birth, but demonstrated slower growth to 1 year of age than infants born before 1970. Growth trajectories converged after 1 year of age. There was no evidence that relative skeletal age, maternal body mass index, or maternal age were associated with growth. They concluded that the recent birth cohorts might be characterized not only by greater birth size, but also by subsequent catch-down growth. And trends over time in human growth do not increased monotonically, and growth velocity in the first year might had declined compared with preceding generations.

Bjelica et al. (2012) analyzed 285 students (178 men and 107 women) from the University of Montenegro to examine the body height in both sexes of Montenegrin adults nowadays. They used mean, standard deviation, *t* test, correlation coefficient and linear regression for the analyzing purpose. Their results showed that male Montenegrins were 183.21±7.06 cm tall and had an arm span of 185.71±8.17 cm, while female Montenegrins were 168.37±5.27 cm tall and had an arm span of 168.13±6.58 cm. Comparing the results with other studies had shown that both sexes of Montenegrins make Montenegro the second tallest nation in the world, while arm span reliably predicts body height in both sexes. However, these estimated equations that had been obtained among the Montenegrins were substantially different than in all other populations, since arm span was close to body height: in males 2.50±4.15 cm more than the body height and in females 0.24±3.88 cm less than the body height.

Chittawatanarat et al. (2012) analyzed 2000 volunteers and were divided consecutively according to both age and gender to develop a formula for height prediction with acceptable validity. They used linear regression with ten parameters model to create a predictive formula. They showed that the demispan, sitting height and knee height were important for the predictive formula. All single parameters and the highest predictive value of double (sitting and knee height) and triple regression models (demispan, sitting and knee height) were proposed and these were modified into a simple formula. The simple formula had more than 90% precision with an error of up to 10 cm in the validation group (89.7 to 99.0% in range). Of these, knee height had the least predictive error in all subgroups. The double and triple models had decreased error only in the younger group. And, they concluded that the anthropometric parameters with demispan, sitting height, knee height and combination could be applied to height prediction in the adult Thai with acceptable error. These formulas should be applied only in people who could not be directly measured.

Pomeroy and Stock (2012) analyzed the regression equations using adult stature and body mass for estimating stature from bone lengths. They proposed new sample-specific regression equations. Anatomical stature reconstruction was further complicated by artificial cranial modification (ACM) influencing cranial height in Andean samples, so this problem was investigated in the current sample. Although ACM had minimal impact here, the possibility should be explored in other samples before anatomical stature estimation was attempted. Recommendations were also made for estimating body mass from femoral head diameter. The mean of three previously published equations was shown to offer minimal bias and the most reliable estimate of body mass in the study samples.

Choksi et al. (2014) analyzed the total number of 500 subjects that includes 200 boys and 300 girls on the young adult population in the age range of 21-25 years to investigate the possible correlation between the palm lengths with the stature of individual. It was possible to deduce the significant correlation coefficient and multiplication factor for estimation of stature from palm length. The multiplication factor so deduced has been applied and regression analysis was done, and was found to be significant and reliable.

Kuninaka and Matsushita (2014) developed a numerical growth model that could predict the statistical properties of the height distribution of Japanese children. Their previous studied had clarified that the height distribution of school children showed transition from the lognormal distribution to the normal distribution during puberty. They demonstrated by simulation that the transition occurs owing to the variability of the onset of puberty.

Rahmandad (2014) discussed the first mechanism-based model spanning full individual life and capturing changes in body weight, composition and height. Integrating previous empirical and modeling findings and validated against several additional empirical studies, the model replicates key trends in human growth including (1) Changes in energy requirements from birth to old ages (2) Short and long-term dynamics of body weight and composition, and (3) Stunted growth with chronic malnutrition and potential for catch up growth. From obesity policy analysis to treating malnutrition and tracking growth trajectories, the model could address diverse policy questions.

2.3 Research Gap

From the above review of literature, some of the researchers predicted final stature through skeletal age, and through the asymptotic curve fitting with longitudinal individual stature only. But, there has hardly been any research which has used the asymptotic curve to model the human stature as a function of several variables. Growth in stature in any living organism depends not only on their age but also body weight, chest circumference, sitting height, genetic factors, maternal illnesses during pregnancy, socio-economic disadvantages during and after pregnancy, social/emotional problems during childhood, poor nutrition and environmental or emotional deprivation and so on. Most of the growth models (Gompertz model, Logistic model, Jenss model, Count model, Double logistic model, PB models, ICP model, Reed models, SSC model, JPPS model, JPA-1 model, JPA-2 model, Modified ICP model, and BTT model) have considered stature or weight as a function of age only. Among them, the JPA-2 model fitted better than all other asymptotic models till 1991 (Jolicoeur et al., 1992). While, BTT model was found to be better than JPA-2 model (Rahman et al., 2004). Like age, it is important to incorporate other possible predictors in the model as they have significant influence on stature. Thus, higher dimensional growth model is necessary for increasing precession of the growth model as well as prediction of the final stature. Theoretically, it is possible to address *n*-dimensional (higher dimensional) growth model (Shahin et al., 2013). But, to check the precession of the model, we need to have longitudinal data set. Presently, in our lab, we have longitudinal data sets of age, weight and stature. Thus, it is possible to check three dimensional growth model, numerically. Age at final stature is known exactly but, in case of weight at final stature, all data are missing that needs to estimate. So, if we estimate first the missing value of weight and then in the second step putting the value of weight in the model as stature = *f*(age, weight), we can estimate the final stature.

2.4 Objectives of the Study

To fulfill the above research gaps, the objectives of the study are to:

i. develop new 3D two-stage growth model, and
ii. develop equations for measuring final stature

CHAPTER THREE

MATERIALS OF THE STUDY

"The Hallmark of Good Science is that is Uses Models and 'Theory' But Never Believes Them." - M. Wilk

CHAPTER 3

MATERIALS OF THE STUDY

3.1 Outline

This chapter includes materials of the study, such as definition of age, stature and weight, and data source with missing value estimation procedure.

3.2 Definition of the Variables Studied

Age: A period of human life, measured by years from birth, usually marked by a certain stage or degree of mental or physical development and involving legal responsibility and capacity.

Stature: Human stature is the distance from the bottom of the feet to the top of the head in a human body, standing erect. Usually the stature is measured by Anthropometer or Stadiometer. The technique used for the new-born child and during the first few years of life differs from the normal technique in that measurements are taken with a special instrument and in a horizontal position since the infants are unable to stand. As has been pointed out, at birth, the average length of a newborn is 20 inches; at 1 year, the average height is about 30 inches; at 2 years, the average height is about 35 inches; and at 3 years, the average height is about 38 inches. After 3 years and until puberty, linear growth continues at a relatively constant rate of 2 inches per year. There are variations of course, according to sex, environment and the ethnic or geographic group concerned.

Weight: The term human body weight is used colloquially and in the biological and medical sciences to refer to a person's mass or weight. Body weight is measured in kilograms, a measure of mass, throughout the world. Body weight is the measurement of weight without items located on the person. Practically though, body weight is measured with clothes on, but without shoes or heavy accessories such as mobile phones and wallets and using manual or digital weighing scales. Excess or reduced body weight is regarded as an indicator of determining a person's health, with body volume measurement providing an extra dimension by calculating the distribution of body weight. The normal weight of the new-born child is between 3000 and 3500 gr. However, in weight as in height, there are also natural variations because of sex, ethnic group, socio-economic status, geographic conditions, etc.

3.3 Data

Our study is theoretical based that need to check with real longitudinal data set. This type of data is unavailable in Bangladesh. Thus, secondary longitudinal data of age, weight and stature of 300 Japanese (180 boys and 120 girls), each between 0 to 20 years old and covering birth-years of 1967 to 1977, have been used. Several universities from the Kanto District of Japan were selected and all students with complete information from several classes of the selected universities were included. I have got this data set from my supervisor who had also got the same data set from his supervisor Late Professor Fumio Ohtsuki, Laboratory of Human Morphology, Graduate School of Science, Tokyo Metropolitan University, Hachiojishi, Tokyo

192-0397, Japan. There are several missing values observed in the data set. The usual procedure of missing value estimate is not applicable in this data set because of longitudinal data (time dependent data). In case of first year, missing value is estimated using the average of all other first year values. The rest of the missing value was estimated by the following process. First, the missing value(s) is(are) filled up by the next observed value(s) to obtain the complete series of data for weight and stature. Second, the complete series was smoothed by resistant smoothing method, 4253H-twice, using Minitab software (Cook and Weisberg, 1982). Third, the smoothing series were arranged according to the order of original series. And, finally, the stature or the weight was plotted against years to have smoothed curve, and the missing value was estimated by the visual initial inspection from the curve.

CHAPTER FOUR

METHODS OF THE STUDY

"The Hallmark of Good Science is that is Uses Models and 'Theory' But Never Believes Them." - M. Wilk

CHAPTER 4

METHODS OF THE STUDY

4.1 Outline

This chapter includes methods of the study such as: descriptive statistics, triphasic generalized logistic human growth (BTT) model, extension of BTT model (proposed model), estimation procedure of the proposed model, finding velocity and acceleration of the proposed model, ARIMA model with Box-Cox transformation, Gompertz growth model, logistic growth model, double logistic growth model, model selection criteria, multiple linear regression model, multicollinearity diagnostic, modified HKB estimator of ridge parameter, forward stepwise regression model, regression diagnostic, least absolute shrinkage and selection operator model, relative bound and absolute bound, *n*-fold cross-validations, cross–validity predicted power, and various software.

4.2 Descriptive Statistics

The descriptive statistics which include the well-known methods such as Mean, Standard Deviation and Correlation Coefficients have been used for measuring the central value, dispersion and degree of association, respectively.

4.3 Triphasic Generalized Logistic Human Growth (BTT) Model

***History of the BTT Model*:** Robertson (1908) proposed that the human growth of organism in general occurs in a number of additive, more-or-less independent phases during the course of development. Generally, the timing and intensity of each phase is assumed genetically programmed in the individual, but they may vary in expression according to environmental conditions. According to Robertson, the sum of the three logistic components can be describing human growth in stature but he cannot test the model because of no suitable data. Bock and Thissen (1976), first applied the concept to individual growth using the case from the Berkeley and Fels growth studies. Their analysis showed that the goodness of fit of the triphasic logistic model was good over the range from one year to maturity. In the Bock-Thissen model, the phases represent early-childhood, middle-childhood, and adolescent growth. A further enhancement of the Bock-Thissen model was suggested by du Toit (1992). He suggested the additional of the 'shape' constants of the positive exponentials to the denominators of the logistic functions to control the model in the region of change-over from early to middle childhood, and to provide some asymmetry of the adolescent component. They found values for these constants that tend to improve the fit of the model, by repeated trials with the Berkeley and Fels data. They refer to this triphasic generalized logistic model as the BTT model.

***Mathematical Explanation of the BTT Model*:** Bock–Thissen–du Toit developed the triple logistic functional model which is properly known as the BTT model. This model was based on the concept that mature size is a summation of three processes says early, middle and adolescent, each of which can be described by a logistic function. Mathematically, the BTT model can be defined as sum of the three generalized logistic terms. The form of the logistic term is:

$$\frac{a}{\left[1+e^{-(bt+c)}\right]^{d}}$$

where, t is the time (age) variable; a, b, c and d are the amount of growth, slope, intercept and fixed shape constant contributed by the term, respectively. The quantity $z = (bt+c)$ in the exponential function is the 'logit'.

Thus, the triphasic generalized logistic model can be written as

$$y=\frac{a_1}{[1+e^{-(b_1t+c_1)}]^{d_1}}+\frac{a_2}{[1+e^{-(b_2t+c_2)}]^{d_2}}+\frac{a_3}{[1+e^{-(b_3t+c_3)}]^{d_3}} \tag{4.1}$$

where, y is the stature at age t; the parameters a_1, a_2 and a_3 decompose the amount of growth of stature contributed respectively by the early, middle and adolescent growth phase; the set of parameters $(a_1,\ b_1,\ c_1)$, $(a_2,\ b_2,\ c_2)$ and $(a_3,\ b_3,\ c_3)$ refer to the parameters of early, middle and adolescent phases of growth, respectively. And, d_1, d_2 and d_3 are fixed shape parameters for early, middle and adolescent phases of growth, respectively.

4.4 Extension of Triphasic Generalized Logistic Human Growth Model

4.4.1 Proposed Model

Total body growth especially stature should be dependent on many factors e.g., age, body weight, chest circumference, sitting height, genetic factors, and maternal illnesses during pregnancy and so on. To the best of our knowledge, all growth models have proposed on age factor only. We have taken here the opportunity to incorporate other factors (for example age and weight) in the BTT growth model. Why we have used BTT model? Because very recently, it is found that BTT growth model perform well than all other parametric growth models **(**Rahman et al., 2004**)**. Now, let us consider a new proposed model, which is the extension of the BTT model. The general mathematical form of proposed model with (p + 1) variables can be written as follows:

$$y = \frac{a_1}{[1+e^{-(a_{11}x_1+\cdots+a_{1p}x_p+c_1)}]^{d_1}} + \frac{a_2}{[1+e^{-(a_{21}x_1+\cdots+a_{2p}x_p+c_2)}]^{d_2}} \frac{a_3}{[1+e^{-(a_{31}x_1+\cdots+a_{3p}x_p+c_3)}]^{d_3}} \qquad (4.2)$$

In case of three variables, the above proposed model (Eq. 4.2) can be written as:

$$y = \frac{a_1}{[1+e^{-(a_{11}t+a_{12}x+c_1)}]^{d_1}} + \frac{a_2}{[1+e^{-(a_{21}t+a_{22}x+c_2)}]^{d_2}} + \frac{a_3}{[1+e^{-(a_{31}t+a_{32}x+c_3)}]^{d_3}} \qquad (4.3)$$

where, y, t and x are the different measurement of body which may be stature, age, and weight. The parameters a_1, a_2 and a_3 decompose the amount of growth of stature contributed by the early, middle and adolescent growth phase, respectively. The set of parameters $(a_1,\ a_{11},\ a_{12},\ c_1)$, $(a_2,\ a_{21},\ a_{22},\ c_2)$ and $(a_3,\ a_{31},\ a_{32},\ c_3)$ refer to the parameters of early, middle and adolescent phases of growth in Euclidian space, respectively. Also, d_1, d_2, and d_3 are fixed shape parameters for early, middle and adolescent phases of growth in Euclidian space, respectively.

Bock and Thissen (1980) imposed a linear restriction on the parameters of the first and second term to remove the over parameterization problem, but du Toit (1992) later found that setting $c_1 = 0$ serves equally well.

4.4.2 Estimation Process of the Proposed Model

There are two cases occur for the estimating proposed model, such as:

1) The regressors are uncorrelated, and
2) The regressors are correlated.

Case (1): The regressors are uncorrelated

When the regressors are uncorrelated, we can estimate the parameters of the model (Eq. 4.3) directly by Bayesian approach because there is no problem of multicollinearity. The method of Bayesian approach is described as follows:

Before estimating the biological parameters of proposed model, we fixed up the shape parameters. The shape parameters can be estimated by trial and error methods. One way is to fix up the shape parameters such that the error is normal that is done by taking different value of shape parameters and fit the model then check normality of error. Remember that, the value of shape parameter of third phase must be greater than the other two phases. Generally, the value of shape parameters of the first two phases is equal. Now, let us consider the proposed nonlinear growth model (Eq. 4.3) is of the form:

$$y = f(X,\Theta) = \sum_{i=1}^{3} \frac{a_i}{[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]^{d_i}} \tag{4.4}$$

where, y is the dependent variable, $X = (t, x)'$ is the vector of independent variables, and $\Theta = (a_1\, a_{11}\, a_{12}\, c_1\, a_2\, a_{21}\, a_{22}\, c_2\, a_3\, a_{31}\, a_{32}\, c_3)'$ is the parameters vector of the model. When observations of y and X are collected for observation i, the equation (4.4) becomes

$$y_i = f\left(X_i, \Theta\right) + \varepsilon_i ;\ i = 1, 2, 3 \ldots n \tag{4.5}$$

where, the model error $\varepsilon_i \sim N(0,\sigma^2)$ distribution. Since the number of parameters is two third of the number of observation hence the estimation by the conventional least squares method for complex growth models is less than the ideal. Even when the number of observations is sufficient for least squares, the parameters may not all be identifiable if the observations are poorly positioned. A much better method for fitting growth models is Bayes model estimation which chooses among a specified population of growth curves.

The random vector of parameters Θ is assumed to follow $N(\mu,\sigma^2)$ distribution in the population. Let us consider a squared error loss function

$$l(\Theta,\hat{\Theta}) = (\Theta - \hat{\Theta})^2$$

Since, $\varepsilon \sim N(0,\sigma^2)$, then

$$E(y\,|\,X,\Theta) = f(X;\Theta) + E(\varepsilon)$$

$$E(y\,|\,X,\Theta) = f(X;\Theta) + 0$$

$\therefore E(y\,|\,X,\Theta) = f(X;\Theta)$, and

$$V(y\,|\,X,\Theta) = \sigma^2$$

So that, $y\,|\left(X,\Theta,\sigma^2\right) \in N\left(f\left(X,\Theta\right),\sigma^2\right)$

We consider a sample $\left(y_1, y_2, \cdots y_n\right)$ of size n from the density $f(y\,|\,X,\Theta,\sigma^2)$. Then the likelihood function is defined as follows:

$$L(y\,|\,X,\Theta,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i - f(X_i;\Theta))^2}$$

$$\therefore L(y\,|X,\Theta,\sigma^2) = \frac{1}{\sqrt[n]{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - f(X_i;\Theta))^2} \tag{4.6}$$

Posterior Bayes estimator of (Θ,σ^2) with respect to prior $g(\Theta,\sigma^2)$ can be written as:

$$\pi(\Theta,\sigma^2\,|y) = \frac{L(y\,|(X,\Theta,\sigma^2))g(\Theta,\sigma^2)}{\int_0^{\infty}\int_{-\infty}^{\infty} L(y\,|(X,\Theta,\sigma^2))g(\Theta,\sigma^2)d\Theta d\sigma^2} \tag{4.7}$$

The denominator in the above posterior distribution (Eq. 4.7) is constant. Thus, the posterior distribution can be represented by the form:

$$\pi(\Theta,\sigma^2\,|y) \propto L(y\,|(X,\Theta,\sigma^2))g(\Theta,\sigma^2)$$

$$\therefore \pi(\Theta,\sigma^2\,|y) \propto \frac{1}{\sqrt[n]{\sigma^2}} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - f(X_i;\Theta))^2} g(\Theta,\sigma^2) \tag{4.8}$$

The most important part in the Bayesian regression analysis is to determine the prior distribution. However, it is very difficult to infer the probability distribution of the regression coefficient in the separated basins. Thus, a short of uniform prior distribution is selected to compute the posterior distribution in the study. Sorensen and Gianola (2002) suggested a sort of uniform distribution using variance:

$$g(\Theta,\sigma^2) \propto \frac{1}{\sigma^2} \tag{4.9}$$

From (Eq. 4.8) and (Eq. 4.9), we can write,

$$\pi(\Theta,\sigma^2\,|y) \propto \frac{1}{\sqrt[n]{\sigma^2}} e^{-\frac{1}{2\sigma^2}} \sum_{i=1}^{n}(y_i - f(X_i;\Theta))^2 \frac{1}{\sigma^2}$$

$$\therefore \pi(\Theta,\sigma^2\,|y) \propto (\sigma^2)^{-(\frac{n}{2}+1)} e^{-\frac{1}{2\sigma^2}} \sum_{i=1}^{n}(y_i - f(X_i;\Theta))^2 \tag{4.10}$$

We know that, for a squared error loss function, the mean of the posterior density is the Bayes estimator of Θ.

Now, the Bayes estimator of Θ under squared error loss function is

$$\hat{\Theta} = E(\pi(\Theta)) = \int_{-\infty}^{\infty} \Theta \pi(\Theta, \sigma^2) d\Theta$$

$$\therefore \hat{\Theta} = \int_{-\infty}^{\infty} \Theta (\sigma^2)^{-(\frac{n}{2}+1)} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - f(X_i;\Theta))^2} d\Theta \tag{4.11}$$

The solution of above equation (Eq. 4.11) can be determined numerically. Fisher-Scoring (Newton-Gauss) method is extremely fast, and nearly as robust as MEAP (Minimum Expected a Posteriori) estimation.

Case (2): The regressors are correlated

Principal components regression is applied in case of correlated regressors. We can write the equation (4.3) as follows:

$$y = \frac{a_1}{\left[1 + e^{-(X\beta_1)}\right]^{d_1}} + \frac{a_2}{\left[1 + e^{-(X\beta_2)}\right]^{d_2}} + \frac{a_3}{\left[1 + e^{-(X\beta_3)}\right]^{d_3}}$$

where, $X = (1, t, x)$, $\beta_1 = (c_1, a_{11}, a_{12})'$, $\beta_2 = (c_2, a_{21}, a_{22})'$ and $\beta_3 = (c_3, a_{31}, a_{32})'$

For *n* data,

$$X = \begin{pmatrix} 1 & t_1 & x_1 \\ 1 & t_2 & x_2 \\ \vdots & \vdots & \vdots \\ 1 & t_n & x_n \end{pmatrix} \quad X'X = \begin{pmatrix} n & \sum_{i=1}^{n} t_i & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} t_i & \sum_{i=1}^{n} t_i^2 & \sum_{i=1}^{n} t_i x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} t_i x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix}$$

Let, Z = XT, $\gamma_i = T'\beta_i$; i=1, 2, 3. Then, $T'X'XT = Z'Z = \Lambda$ and $\Lambda = \text{diag}(\lambda_0, \lambda_1, \lambda_2)$ is a 3×3 diagonal matrix of the eigenvalues of $X'X$ and T is a 3×3 orthogonal matrix whose columns are the eigen vectors associated with λ_0, λ_1, λ_2. We can define a new set of orthogonal regressors, such as $Z = (Z_0, Z_1, Z_2)$ which is the column of Z are referred to as principal components.

Therefore,

$$y = \frac{a_1}{\left[1+e^{-(Z\gamma_1)}\right]^{d_1}} + \frac{a_2}{\left[1+e^{-(Z\gamma_2)}\right]^{d_2}} + \frac{a_3}{\left[1+e^{-(Z\gamma_3)}\right]^{d_3}}$$

The principal components regression approach used less than full set of principal components to combat multicollinearity in the model. In principal components estimators, we assume that the regressors are arranged in order of decreasing eigenvalues:

$$\lambda_0 \geq \lambda_1 \geq \lambda_2 > 0$$

Let us suppose that the last s of these eigenvalues approximately equal to zero. In principal components regression the principal components corresponding to near-zero eigenvalues are removed from the analysis and Bayesian estimate defined in case (1) section is applied to the remaining components. That is,

$\hat{\gamma}_{ipc} = B\hat{\gamma}_i$; i = 1, 2, 3 and the fitted model can be written as:

$$\hat{y} = \frac{a_1}{\left[1+e^{-(Z\hat{\gamma}_1)}\right]^{d_1}} + \frac{a_2}{\left[1+e^{-(Z\hat{\gamma}_2)}\right]^{d_2}} + \frac{a_3}{\left[1+e^{-(Z\hat{\gamma}_3)}\right]^{d_3}}$$

Replacing Z by the linear combination of X, we get

$$\hat{y} = \frac{a_1}{\left[1+e^{-(XT\hat{\gamma}_1)}\right]^{d_1}} + \frac{a_2}{\left[1+e^{-(XT\hat{\gamma}_2)}\right]^{d_2}} + \frac{a_3}{\left[1+e^{-(XT\hat{\gamma}_3)}\right]^{d_3}}$$

$$\hat{y} = \frac{a_1}{\left[1+e^{-\left(X\hat{\beta}_{1pc}\right)}\right]^{d_1}} + \frac{a_2}{\left[1+e^{-\left(X\hat{\beta}_{2pc}\right)}\right]^{d_2}} + \frac{a_3}{\left[1+e^{-\left(X\hat{\beta}_{3pc}\right)}\right]^{d_3}}$$

Thus, the principal component estimator can be written as follows:

$\hat{\beta}_{ipc} = T\hat{\gamma}_{ipc}$; $\underline{i}$=1, 2, 3 [since $\hat{\gamma}_i = T'\beta_i$ i=1, 2, 3 and $T'T = I_3$]

4.4.3 Methods of Finding Velocity and Acceleration

Velocity: Velocity is a term used for a rate of change. That is, velocity is defined as the ratio of the directed displacement Δr (say) to the required time Δt (say). That is,

Velocity $= \dfrac{\Delta r}{\Delta t}$

Again, let us consider the proposed model as

$$y = f(X,\Theta) = \sum_{i=1}^{3} \frac{a_i}{[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]^{d_i}}$$

Velocity for the variable *t*, when *x* is constant

$$\frac{\partial y}{\partial t} = \sum_{i=1}^{3} \frac{a_i d_i a_{i1} e^{-(a_{i1}t+a_{i2}x+c_i)}}{[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]^{d_i}[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]}$$

Velocity for the variable *x*, when *t* is constant

$$\frac{\partial y}{\partial x} = \sum_{i=1}^{3} \frac{a_i d_i a_{i2} e^{-(a_{i1}t+a_{i2}x+c_i)}}{[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]^{d_i}[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]}$$

Acceleration: Acceleration of a function $y = f(X,\Theta)$ is defined as the second derivative of *y* with respect to *x*. That is, Acceleration = $\dfrac{\partial^2 y}{\partial x^2}$

Acceleration for the variable *t*, when *x* is constant

$$\frac{\partial^2 y}{\partial t \partial t} = \sum_{i=1}^{3} \left[\frac{a_i d_i^2 a_{i1}^2 [e^{-(a_{i1}t+a_{i2}x+c_i)}]^2}{[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]^{d_i+2}} - \frac{a_i d_i a_{i1}^2 e^{-(a_{i1}t+a_{i2}x+c_i)}}{[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]^{d_i+1}} + \frac{a_i d_i a_{i1}^2 [e^{-(a_{i1}t+a_{i2}x+c_i)}]^2}{[1+e^{-(a_{i1}t+a_{i2}x+c_i)}]^{d_i+2}} \right]$$

Acceleration for the variable *x*, when *t* is constant

$$\frac{\partial^2 y}{\partial x \partial x} = \sum_{i=1}^{3} \left[\frac{a_i d_i^2 a_{i2}^2 [e^{-(a_{i1}t + a_{i2}x + c_i)}]^2}{[1 + e^{-(a_{i1}t + a_{i2}x + c_i)}]^{d_i + 2}} - \frac{a_i d_i a_{i2}^2 e^{-(a_{i1}t + a_{i2}x + c_i)}}{[1 + e^{-(a_{i1}t + a_{i2}x + c_i)}]^{d_i + 1}} + \frac{a_i d_i a_{i2}^2 [e^{-(a_{i1}t + a_{i2}x + c_i)}]^2}{[1 + e^{-(a_{i1}t + a_{i2}x + c_i)}]^{d_i + 2}} \right]$$

To compute the Hessian matrix, we need various derivative of $y = f(X, \Theta)$ with respect to *t* and, *x* as follows:

Now, Partial derivative of y with respective the variable *t* and *x*

$$\frac{\partial^2 y}{\partial t \partial x} = \sum_{i=1}^{3} \left[\frac{a_i d_i^2 a_{i1} a_{i2} [e^{-(a_{i1}t + a_{i2}x + c_i)}]^2}{[1 + e^{-(a_{i1}t + a_{i2}x + c_i)}]^{d_i + 2}} - \frac{a_i d_i a_{i1} a_{i2} e^{-(a_{i1}t + a_{i2}x + c_i)}}{[1 + e^{-(a_{i1}t + a_{i2}x + c_i)}]^{d_i + 1}} + \frac{a_i d_i a_{i1} a_{i2} [e^{-(a_{i1}t + a_{i2}x + c_i)}]^2}{[1 + e^{-(a_{i1}t + a_{i2}x + c_i)}]^{d_i + 2}} \right]$$

4.4.4 Methods of Finding Minimum and Maximum Values

To compute the minimum and maximum values of the proposed model (Eq. 4.3), the Gradient vector and Hessian matrix can be used. The proposed model can be written as the following form:

$$y = f(t, x) = \sum_{i=1}^{3} \frac{a_i}{[1 + e^{-(a_{i1}t + a_{i2}x + c_i)}]^{d_i}}$$

The Gradient vector and Hessian matrix of a function of two variables f (t, x) are defined respectively as follows:

$$\nabla f(t, x) = \begin{pmatrix} \dfrac{\partial f(t, x)}{\partial t} \\ \dfrac{\partial f(t, x)}{\partial x} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f(t, x)}{\partial t} & \dfrac{\partial f(t, x)}{\partial x} \end{pmatrix}' \text{ and } \nabla^2 f(t, x) = \begin{pmatrix} \dfrac{\partial^2 f(t, x)}{\partial t \partial t} & \dfrac{\partial^2 f(t, x)}{\partial t \partial x} \\ \dfrac{\partial^2 f(t, x)}{\partial x \partial t} & \dfrac{\partial^2 f(t, x)}{\partial x \partial x} \end{pmatrix}$$

Let us consider any value of t and x are t^* and x^*, respectively such that $\nabla f(t^*, x^*) = 0$ and satisfy any one of the following condition:

Case 1: $f(t^*, x^*)$ have a local minimum value if

$$\frac{\partial^2 f(t^*, x^*)}{\partial t \partial t} > 0, \begin{vmatrix} \dfrac{\partial^2 f(t^*, x^*)}{\partial t \partial t} & \dfrac{\partial^2 f(t^*, x^*)}{\partial t \partial x} \\ \dfrac{\partial^2 f(t^*, x^*)}{\partial x \partial t} & \dfrac{\partial^2 f(t^*, x^*)}{\partial x \partial x} \end{vmatrix} > 0$$

Case 2: $f(t^*, x^*)$ have a local maximum value if

$$\frac{\partial^2 f(t^*, x^*)}{\partial t \partial t} < 0, \begin{vmatrix} \dfrac{\partial^2 f(t^*, x^*)}{\partial t \partial t} & \dfrac{\partial^2 f(t^*, x^*)}{\partial t \partial x} \\ \dfrac{\partial^2 f(t^*, x^*)}{\partial x \partial t} & \dfrac{\partial^2 f(t^*, x^*)}{\partial x \partial x} \end{vmatrix} > 0$$

Case 3: There is no local maxima or minima if $f(t^*, x^*)$ does not satisfy the two conditions described above, and which is called a saddle point.

4.5 Model Selection Criteria

Model selection criteria are very important. There are different algorithms for selecting best model; however, the selection is difficult. Selection of model is data dependent. For this, mean squared error (MSE) and root mean squared error (RMSE) may be used. The mathematical formula of MSE and RMSE are as follows:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2 \quad \text{and} \quad RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(Y_t - \hat{Y}_t\right)^2}$$

The MSE can be used to test the performance of the existence triphasic generalized logistic human growth (BTT) model and the proposed extension of triphasic generalized logistic human growth (BTT) model.

4.6 Predictions of Final Stature

Suppose final stature attains at age 25. So, predicting the final stature, weight at age 25 is needed which can be estimated in two stages.

4.6.1 First Stage

In this stage, the weight at final stature is estimated by one of the appropriate method, namely (i) autoregressive integrated moving average (ARIMA) model with Box−Cox transformation, (ii) Gompertz, (iii) logistic and (iv) double logistic models. Smallest mean square error (section 4.5) is taken as the model selection criteria. The statistical software R i386 3.0.1 with the package "*forecast*" is used to automatically chosen the optimum order of ARIMA model and the optimum value of Box−Cox parameter (λ) for each data set. The remaining models are estimated by the STATISTICA 8.0 software. The description and mathematical form of these models are given bellow:

Autoregressive Integrated Moving Average (ARIMA) Model with Box-Cox Transformation: Box and Cox (1964) have proposed a family of transformations that can be used with non-negative responses with transformations in common use, including reciprocals, logarithms and square roots. Transformations of data designated to achieve a specified purpose, e.g., stability of variance, additively of effects and symmetry of the density. If one is successful in finding a suitable transformation, the ordinary method for analysis will be available. Among the many parametric transformations, the Box-Cox family is commonly utilized.

Let y_t be a random variable on the positive half-line. Then the Box–Cox transformation of y_t with power parameter λ is defined as:

$$y_t^{\lambda} = \begin{cases} \dfrac{y_t^{\lambda} - 1}{\lambda} & \text{for} \lambda \neq 0 \\ \log y_t & \text{for} \lambda = 0 \end{cases}$$

The formula $\frac{Y_t{}^{\lambda} - 1}{\lambda}$ will be chosen so that y_t^{λ} is continuous as λ tends to zero and monotonic increasing with respect to y_t for any λ. The power parameter λ is estimated by a graphical technique or by the maximum-likelihood method. Unfortunately, a closed form for the estimator λ can be rarely found. Hence, the plot of the maximum likelihood against λ will give the value of $\hat{\lambda}$ to have fit the transformed data for fitting any model.

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. The ARIMA model is applied in some cases where data show evidence of non-stationarity. The model is generally referred to as an ARIMA(*p*, *d*, *q*) model, where *p*, *d*, and *q* are integers greater than or equal to zero and refer to the order of the autoregressive, integrated, and moving average parts of the model, respectively. The ARIMA(*p*, *d*, *q*) model can be written in a compact way as follows:

$$\phi(B)\nabla^{d} y_t = c + \theta(B)\varepsilon_t$$

where, *WN* stands for white noise; *B* means backshift operator; $\nabla^{d} = (1 - B)^{d}$ (The *d* order differencing operator); $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_P B^p$ (The *p* order of AR operator); $\theta(B)=1+\theta_1 B+\theta_2 B^2+\cdots+\theta_q B^q$ (The q order of MA operator); ε_t is random shocks distributed normally with mean zero and constant variance σ^2; c is drift (i.e., constant) and y_t is any Box-Cox transformed time series.

Gompertz Growth Model: The Gompertz growth model can be used to measure the weight at final stature. The mathematical form and description of parameters have been explained in chapter 2 in review of literature section.

Logistic Growth Model: The logistic growth model can be used to measure the weight at final stature. The mathematical form and description of parameters have been explained in chapter 2 in review of literature section.

Double Logistic Growth Model: The double logistic growth model can be used to measure the weight at final stature. The mathematical form and description of parameters have been explained in chapter 2 in review of literature section.

4.6.2 Second Stage

In this stage, the final stature is predicted by the proposed model (Eq. 4.3). The regressand and regressors variable of the proposed model are stature; and age and weight, respectively. Using the data of stature, age and weight for each data set; the proposed model is estimated and the corresponding estimated values of all parameters were obtained. Suppose, the estimated proposed model for i^{th} (i = 1 to 180 for boys and i = 1 to 120 for girls) data set was as follows:

$$\hat{y}_{25,i}=\frac{\hat{a}_{1,i}}{[1+e^{-(\hat{a}_{11,i}t_{25,i}+\hat{a}_{12,i}x_{25,i}+\hat{c}_{1,i})}]^{d_1}}+\frac{\hat{a}_{2,i}}{[1+e^{-(\hat{a}_{21,i}t_{25,i}+\hat{a}_{22,i}x_{25,i}+\hat{c}_{2,i})}]^{d_2}}+\frac{\hat{a}_{3,i}}{[1+e^{-(\hat{a}_{31,i}t_{25,i}+\hat{a}_{32,i}x_{25,i}+\hat{c}_{3,i})}]^{d_3}}$$

where, $\hat{y}_{25,i}$, $t_{25,i}$ and $x_{25,i}$ represented the estimated final stature, age at final stature and weight at age 25 for i^{th} data set, respectively. The estimated parameters of the proposed model were denoted by $\hat{a}_{1,i}$, $\hat{a}_{11,i}$, $\hat{a}_{12,i}$, $\hat{c}_{1,i}$, $\hat{a}_{2,i}$, $\hat{a}_{21,i}$, $\hat{a}_{22,i}$, $\hat{c}_{2,i}$, $\hat{a}_{3,i}$, $\hat{a}_{31,i}$, $\hat{a}_{32,i}$ and $\hat{c}_{3,i}$ for i^{th} data set. And, d_1, d_2 and d_3 were known fixed shape parameters. Using the estimated proposed model for each data set with corresponding age equal 25 and estimated weight at age 25 (from first stage), the value of estimated final stature can be obtained.

4.7 Building Equations for Predicting Final Stature

Suppose, we are interested to estimate the final stature using the earlier statures from fitted 3D model (Eq. 4.3) and parent statures. This can be done using forward stepwise ridge regression model and LASSO model. Before hand, we have to check the following steps.

Linearity of Final Stature on Other Regressor: Stature at different ages can be extracted from fitted distance curve (Eq. 4.3). To check the linearity of final stature on earlier stature at different ages and parent stature, correlation matrix plot can be applied as in Figure 4.1a and 4.1b. Figure 4.1a and Figure 4.1b show linear relationships between the final stature and the stature at the ages 2–13 and parent statures for boys and girls, respectively. Thus, linear regression method can be applied.

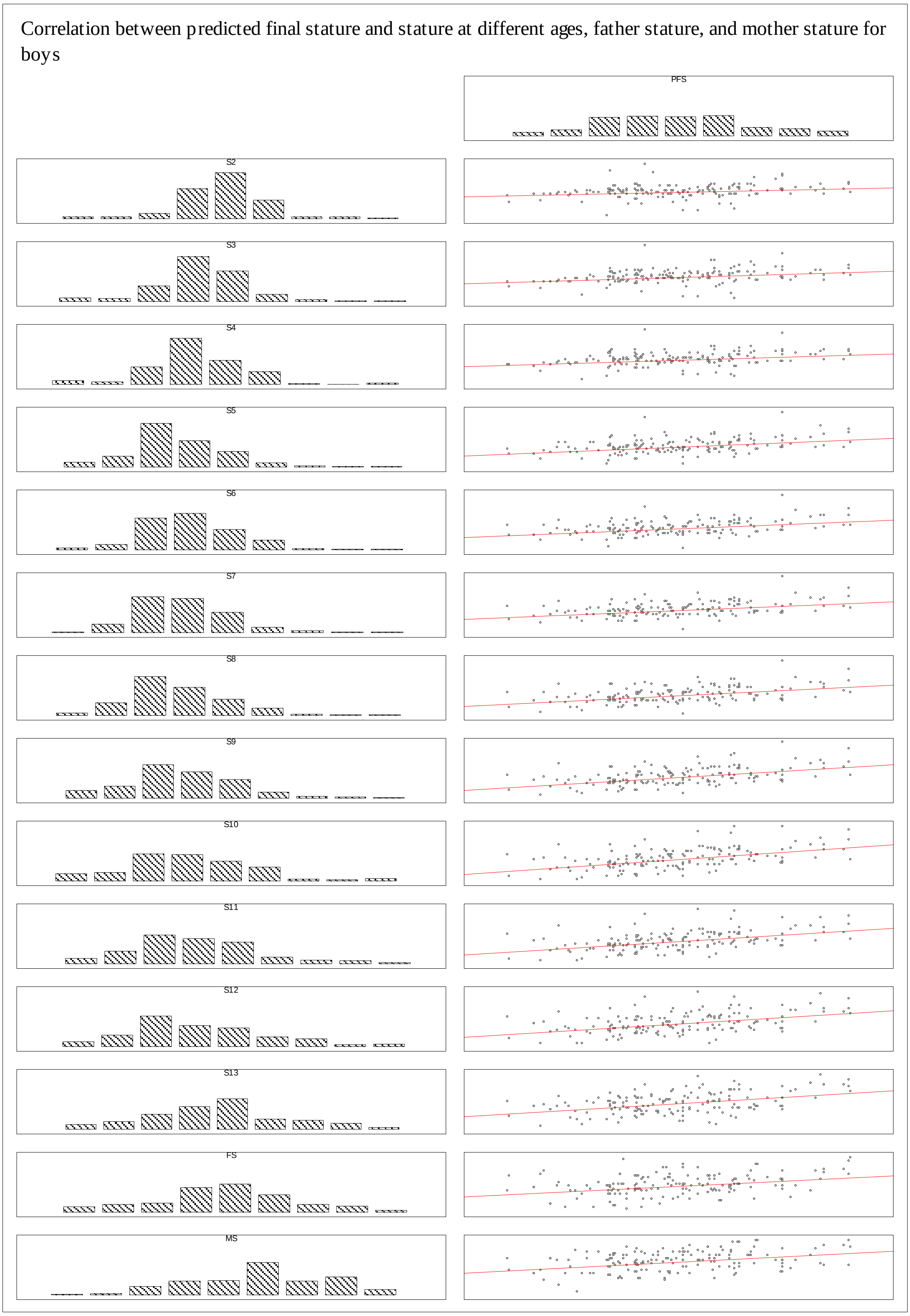


Figure 4.1a Correlation matrix plot between predicted final stature (PFS) and stature at different ages, father stature and mother stature for boys. The statures are drawn from the estimated distance curves of the proposed model. The father stature (FS) and mother stature (MS) come from the individual recode. Sj is the predicted stature at age j, j = 2, 3,…,13. In every element in the matrix, the *X*-axis is for PFS. Only the pattern, either linear or nonlinear, is of concern. No scales are shown here.

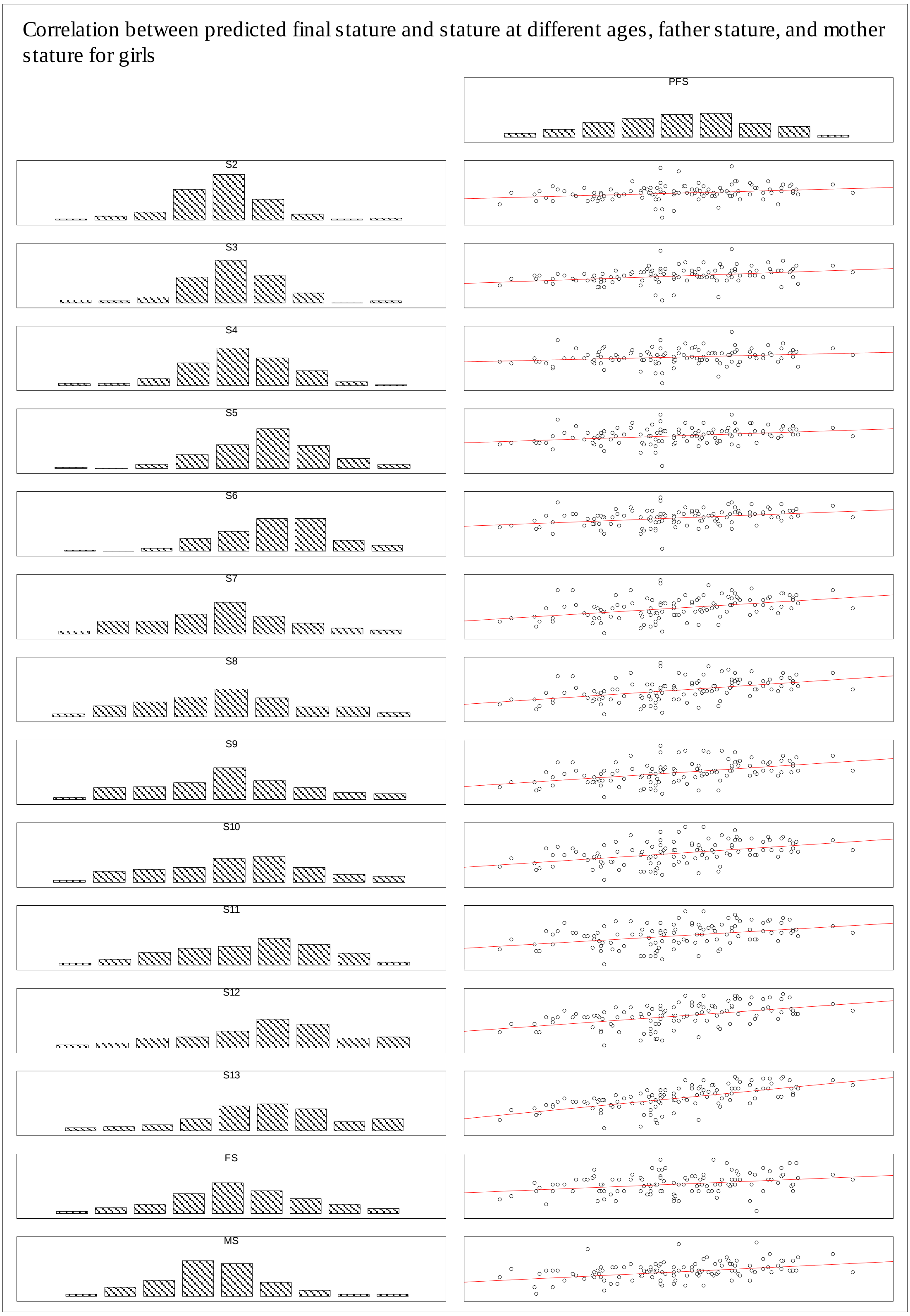


Figure 4.1b Correlation matrix plot between predicted final stature (PFS) and stature at different ages, father stature and mother stature for girls. The statures are drawn from the estimated distance curves of the proposed model. The father stature (FS) and mother stature (MS) come from the individual recode. Sj is the predicted stature at age j, j = 2, 3, …, 13. In every element in the matrix, the *X*-axis is for PFS. Only the pattern, either linear or nonlinear, is of concern. No scales are shown here.

Multiple Linear Regression Model: We can consider the multiple linear regressions model of *PFS* on stature at each age from 2 to 13 years and parent statures as their relationships are linear. A regression equation without an intercept is applicable here. Let us suppose that the stature at the age j, j = 2, 3, …, 13, father stature, mother stature and predicted final stature are denoted by S_j, j = 2, 3, …, 13, *FS, MS* and *PFS*, respectively. Therefore, the multiple linear regression models without an intercept term can be expressed mathematically as follows:

$$PFS_i = \beta_2 S_{i2} + \beta_3 S_{i3} + \cdots + \beta_{13} S_{i13} + \beta_{14} FS_i + \beta_{15} MS_i + \varepsilon_i; \;\; i = 1, 2, \cdots, n$$

where, β_j, j = 2, 3, …, 13, β_{14} and β_{15} are the partial regression coefficients and n is the sample size. The variable *PFS* as regressand and the variables $S_j, j = 2, 3, \ldots 13,$ FS and MS are the regressors and ε is error term assumed to be distributed as normally with mean zero and constant variance.

The regression model in term of the observations, may be written in matrix notation as

$$PFS = S\beta + \varepsilon$$

where,

$$PFS = \begin{pmatrix} PFS_1 \\ PFS_2 \\ \vdots \\ PFS_n \end{pmatrix}, \quad S = \begin{pmatrix} S_{12} & \cdots & S_{113} & FS_1 & MS_1 \\ S_{22} & \cdots & S_{213} & FS_2 & MS_2 \\ \vdots & \cdots & \vdots & \vdots & \vdots \\ S_{n2} & \cdots & S_{n13} & FS_n & MS_n \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_2 \\ \beta_3 \\ \vdots \\ \beta_{15} \end{pmatrix} \text{ and } \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

In general, *PFS* is an (n×1) vector of the observations; *S* is an (n×14) matrix generated by regressor variables, β is a (14×1) vector of the regression coefficients, and ε is an (n×1) vector of random errors.

Thus the ordinary least squares estimator of β is

$$\hat{\beta} = \left(S^T S\right)^{-1} S^T PFS \tag{4.12}$$

There may be possibility that the model to have the problem of multicollinearity which needs to diagnose.

Multicollinearity Diagnostic: Tolerance and Variance Inflection Function (VIF) can be used to measure the problems of multicollinearity. The diagonal elements of the $C = (S'S)^{-1}$matrix are very useful in detecting multicollinearity. Let C_{kk} be the k th diagonal element of *C*, and can be defined as $C_{kk} = (1 - R_k^2)^{-1}$, where, R_k^2 is the coefficient of the determination obtained when S_k is regressed on the remaining 13 regressors. The tolerance is defined as inverse of C_{kk}. The closer is tolerance to zero, the greater the degree of collinearity of that variable with the other regressors. On the other hand, the closer tolerance is to 1, the greater the evidence that S_k is not collinear with the other regressors. Marquardt (1970) has called $C_{kk} = (1 - R_k^2)^{-1}$ the "variance inflation factor" (VIF). The VIF for each term in the model measures the combined effect of the dependencies among the regressors on the variance of that term. One or more large VIFs indicate multicollinearity. Practical experience indicates that if any of the VIFs equal to 1, then the regressors are not correlated, if VIFs is 1 to 5, then the regressors are moderately correlated, and VIFs exceeds 5, then the regressors are highly correlated. The tolerance and VIF (Table 5.6 in Chapter 5) clearly indicates the multicollinearity problem exists in the data set. Therefore, special types of regression method (namely stepwise, ridge regression, LASSO etc.) can be applied to reduce the problems of multicollinearity and select appropriate important variables.

Forward Stepwise Ridge Regression Model: Since the multicollinearity problems exist in the data set, hence the least squares produce very poor estimates of the regression coefficients because of the data are non-orthogonal. The variance of the least squares estimates of the regression coefficients may be considerably inflated, and the length of the vector of least squares parameter estimates is too long on the average. This implies that the absolute value of the least squares estimates are too large, and that they are very unstable. The least square estimate of the multiple linear regression model (eq. 4.12) is as follows:

$$\hat{\beta} = \left(S^T S\right)^{-1} S^T PFS$$

The ridge estimator is found by solving a slightly modified version of the above equation. Specifically, we define the ridge estimator $\hat{\beta}_R$ as

$$\hat{\beta}_R = \left(S^T S + kI\right)^{-1} S^T PFS$$

where, $k \geq 0$ is a constant selected by the analyst. The parameter, namely k *is* called ridge parameter. The value of k is estimated by maximum likelihood estimator proposed by Hoerl, Kennard and Baldwin (HKB estimator) (Hoerl et al., 1975) which is properly known as Modified HKB Estimator of ridge parameter.

The stepwise regression can be used to find out the 'best' set of explanatory variables for a regression model. In this method, forward stepwise regression proceeds by introducing the regressor variables one at a time. Forward stepwise ridge regression to individually add the independent variables from the model at each step with including the value of ridge parameter of the regression (depending on your choice of F to enter) until the 'best' regression model is obtained. The F to enter value determines how significant the contribution of a variable in the regression equation has to be in order for it to be added to the equation.

Least Absolute Shrinkage and Selection Operator (LASSO): A "LASSO" is usually recognized as a loop of rope that is designed to be thrown around a target and tighten when pulled. Least absolute shrinkage and selection operator is fittingly being used as a metaphor of l_1 constraint applied to linear model. The ridge regression model can be used for the collinearity data. But, the ridge regression penalty ($\sum \beta_j^2$), although it helps with obtaining less variable estimates, has two big shortcomings in this setting: (1) Heavy bias toward zero for large regression coefficients (2) Interpretability: unimportant coefficients may be shrunken towards zero, but they're still in the model. We know that if many features are correlated, least absolute shrinkage and selection operator (LASSO) will just pick one.

Consider the usual linear regression model with data (S_{ij}, j = 2, 3, …13, FS_i, MS_i, PFS_i), i =1,…,n, where PFS_i is the response variable of the i^{th} observation and all other variables are regressor. The Ordinary Least Squares (OLS) regression method finds the linear combination of the S_{ij}, j = 2, 3, …13, FS_i and MS_i (i =1,…,n) that minimizes the residual sum of squares. However, if number of regressors is large or the regression coefficients are highly correlated (multicolinear), the OLS may yield estimates with large variance which reduces the accuracy of the prediction. A widely-known method to solve this problem is the ridge regression and the method of selecting subset. As an alternative to these techniques, Tibshirani (1996) presented "LASSO" which minimized the residual sum of squares subject to the sum of absolute values of the coefficient being less than a constant.

$$\hat{\beta}^{L} = \arg\min\left[\sum_{i=1}^{n}\left\{PFS_i - \left(\sum_{j=2}^{13}\beta_j S_{ij} + \beta_{14}FS_i + \beta_{15}MS_i\right)\right\}^2\right]$$

subject to

$$\sum_{j=2}^{15}\left|\hat{\beta}_j^{L}\right| \le t \quad (\text{constant})$$

if $t > \sum_{j=2}^{15}\left|\hat{\beta}_j^{0}\right|$, then the LASSO algorithm will yield the same estimate as OLS estimate. However, if $0 < t < \sum_{j=2}^{15}\left|\hat{\beta}_j^{0}\right|$, then the problem is equivalent to

$$\hat{\beta}^{L} = \arg\min\left[\sum_{i=1}^{n}\left\{PFS_i - \left(\sum_{j=2}^{13}\beta_j S_{ij} + \beta_{14}FS_i + \beta_{15}MS_i\right)\right\}^2 + \lambda\sum_{j=2}^{15}\left|\beta_j\right|\right]$$

$\lambda > 0$, It will be shown later that the relation between λ and LASSO parameter t is one-to-one. Due to the nature of the constraint, LASSO tends to produce some coefficients to be exactly zero. Compared to the OLS, whose predicted coefficient $\hat{\beta}^0$ is an unbiased estimator of β, both ridge regression and LASSO sacrifice a little bias to reduce the variance of the predicted values and improve the overall prediction accuracy.

The tuning parameter $\sum_{j=2}^{15}\left|\hat{\beta}_j^{L}\right| = t$ is called LASSO parameter, which is also recognized as the absolute bound. Here we define another parameter, s, as the relative bound.

$$s = \frac{\sum_{i=2}^{15}\left|\hat{\beta}_j^{L}\right|}{\sum_{i=2}^{15}\left|\hat{\beta}_j^{o}\right|}, \quad s \in [0, 1]$$

The relative bound can be seen as a normalized version of LASSO parameter. There are two algorithms mentioned in Tibshirani (1996) to compute the best *s:* (i) *n*-fold cross-validation and (ii) generalized cross-validation (GCV).

Cross-validation is a general procedure that can be applied to estimate tuning parameters in a wide variety of problems. The bias in RSS is a result of using the same data for model fitting and model evaluation. Cross validity can reduce the bias of RSS by splitting the whole data into two subsamples: a training (calibration) sample for model fitting and a test (validation) sample for model evaluation. The idea behind the cross-validation is to recycle data by switching the roles of training and test samples.

The optimal s can be denoted by $\hat{s}$. Prediction error can be estimated for the LASSO procedure by ten-fold cross-validation (Tibshirani, 1996). The LASSO is indexed in terms of s, and the prediction error is estimated over a grid of values of s from 0 to 1 inclusive. We wish to predict with small variance, thus we wish to choose the constraint s as small as we can. The value $\hat{s}$ which achieves the minimum predicted error is selected (Tibshirani, 1996).

4.8 Regression Diagnostic

Coefficient of Determination: In case of the two-variables, the square of simple correlation coefficient (r^2) measures the goodness of fit of the regression equation; that is, it gives the proportion of the total variation in the dependent variable explained by the (single) explanatory variable. This notation of r^2 can be easily extended to regression models containing more than two variables. Thus, in the three or more variable models we would like to know the proportion of the variation in dependent variable explained by the regressor variables jointly. The quantity that gives this information is known as the multiple coefficient of determination and is denoted by R^2.

Adjusted Coefficient of Determination: An important property of R^2 is that it is a non-decreasing function of the number of explanatory variables or regressors present in the model; as the number of regressors increases, R^2 almost invariably increases and never decreases. In view of this, in comparing two regression models with the same dependent variable but differing number of regressor variables, one should be very wary of choosing the model with the highest R^2. To compare two R^2 terms, one must take into account the number of regressor variables present in the model. This can be done readily if we consider an alternative coefficient of determination, is known as the adjusted R^2, denoted by $\bar{R}^2$. The term adjusted means adjusted for the degree of freedom associated with the sums of squares entering intoR^2.

Zero Intercept: Generally, a regression model has unknown an intercept and the regression coefficients terms and these terms have special meaning. The intercept and regression coefficients terms can be estimated by using sample data set. Sometimes, it is necessary to specify a regression equation without an intercept (intercept forced to zero, regression through the origin). In such a situation, the value of R^2 is higher than the value of R^2 in case of inclusion of an intercept term.

Checking Outliers and Influential Data Points: Generally, an observation is called outlier when it is lie outside of 3 of standardized residuals. Cook (1977) proposed an estimator which is commonly used to estimate of the influence of a data point when performing least squares regression analysis. In a practical ordinary least squares analysis, Cook's distance can be used in several ways: (i) to indicate data points that are particularly worth checking for validity and (ii) to indicate regions of the design space where it would be good to be able to obtain more data points. The Cook's distance is denoted by *D* and it is calculated as:

$$D_i = \frac{\sum_{j=1}^{n} \left(\hat{Y}_j - \hat{Y}_{j(i)} \right)^2}{p \times MSE}$$

The following are the algebraically equivalent expressions (in case of simple linear regression):

$$D_i = \frac{e_i^2}{p \times MSE} \left\{ \frac{h_{ii}}{(1 - h_{ii})^2} \right\}$$

$$D_i = \frac{\left\{ \hat{\beta} - \hat{\beta}^{(-i)} \right\}^T S^T S \left\{ \hat{\beta} - \hat{\beta}^{(-i)} \right\}}{(1 + p) \times s^2}$$

In the above equations: $\widehat{Y}_j$ is the prediction from the full regression model for observation *j*; $\widehat{Y}_{j(i)}$ is the prediction for observation *j* from a refitted regression model in which observation *i* has been omitted; h_{ii} is the *i*th diagonal element of the hat matrix $(S^T S)^{-1} S^T$; e_i is the crude residual (i.e., the difference between the observed value and the value fitted by the proposed model); *MSE* is the mean square error of the regression model; and *p* is the number of fitted parameters in the model.

Remedies for Outliers: Particularly with small *n* (less than 100) multiple regression estimates are not very stable. In other words, single extreme observations can greatly influence the final estimates. Therefore, it is and visible always to use formal statistical procedures to identify outliers and to repeat the analysis after omitting any outliers. Another alternative is to use robust techniques.

Checking Normality of Residuals: The normal probability plot is useful for determining how well a specific theoretical distribution fits the observed data. In the *P-P* plot, the observed cumulative distribution function is plotted against a theoretical cumulative distribution function in order to assess the fit of the theoretical distribution to the observed data. If all points in this plot fall onto a diagonal line (with intercept 0 and slope 1), then we can conclude that the theoretical cumulative distribution approximates the observed distribution well. If the data points do not all fall on the diagonal line, then we can use this plot to visually assess where the data do and do not follow the distribution.

Cross–Validity Predicted Power: Cross validity predictive power (Stevens, 1996, pp-100) is used for testing the validity and stability of the fitted models. The fitted equations with high value of cross validity predicted power is assumed a better representation of the population. The cross validity predictive power, denoted by ρ_{cv}^2, is defined as:

$$\rho_{cv}^2 = 1 - \frac{(N-1)(N-2)(N+1)}{N(N-P-1)(N-P-2)}\left(1-R^2\right)$$

where, N is the sample size, P is the number of predictors in the regression equation and the cross validated R is the correlation coefficient between observed and predicted values of the dependent variable. Using the above statistic, it can be concluded that if the prediction equation is applied to many other samples from the same population, then ($\rho_{cv}^2 \times 100$) % of the variance on the predicted variable would be explained by the regression equation (Stevens, 1996; pp-100).

4.9 Software

The popular software STATISTICA 8.0, Minitab 12.1, Statistical free software R i386 3.0.1 with packages 'forecast', 'MASS', 'lars' and 'lasso2', MS-Excel and MS-Word have been used to the thesis, according to the objective.

4.10 R Code

Let us consider the all data files are located in D drive.

```
#Read the boys data file, namely 'Boys_weight.csv' from D drive

BoysWeight<- read.csv(file = "D:Boys_weight.csv", sep = ",", header = TRUE)

#Select the appropriate value of Box-Cox lambda, order of ARIMA model and forecast value for boys

library(forecast)

lambda<- BoxCox.lambda(BoysWeight, lower = -3, upper = 3)

Model_ARIMA_Boys<- auto.arima(BoysWeight, lambda)

Forecast_Weight_Boys<- forecast(Model_ARIMA_Boys, (25 - length(BoysWeight)))
```

#Read the girls data file, namely 'Girls_weight.csv' from D drive

```
GirlsWeight<- read.csv(file = "D:Girls_weight.csv", sep =
",", header = TRUE)
```

#Select the appropriate value of Box-Cox lambda with ARIMA model and forecast value for girls

```
library(forecast)

lambda<- BoxCox.lambda(GirlsWeight, lower = -3, upper = 3)

Model_ARIMA_Girls<- auto.arima(GirlsWeight, lambda)

Forecast_Weight_Girls<- forecast(Model_ARIMA_Girls, (25 -
length(GirlsWeight)))
```

#Read the boys data file, namely 'Boys_stature.csv' from D drive

```
BoysStature<- read.csv(file = "D:Boys_stature.csv", sep =
",", header = TRUE);
```

#Automatic selects the values of ridge regression parameters for boys

```
library(MASS)

ModelBoys<-lm.ridge(PFS ~ 0 + S2 + S3 + S4 + S5 + S6 + S7 +
S8 + S9 + S10 + S11 + S12 + S13 + FS + MS, data =
BoysStature, lambda = seq(0,0.1,0.001))

select(ModelBoys)
```

#Read the girls data file, namely 'Girls_stature.csv' from D drive

```
GirlsStature<- read.csv(file = "D:Girls_stature.csv", sep = ",", header = TRUE);
```

#Automatic selects the values of ridge regression parameters for girls

```
library(MASS)

ModelGirls<- lm.ridge(PFS ~ 0 + S2 + S3 + S4 + S5 + S6 + S7 + S8 + S9 + S10 + S11 + S12 + S13 + FS + MS, data = GirlsStature, lambda = seq(0,0.1,0.001))

select(ModelGirls)
```

#Read the boys data file, namely 'Boys_data.csv' from D drive

```
BoysData<- read.csv(file = "D:Boys_data.csv", sep = ",", header = TRUE)

yb<- as.numeric(BoysData[,15])

xb<- as.matrix(BoysData[,1:14])
```

#Estimate the parameters of LASSO regression model for boys

```
library(lars)

lasso_fit_Boys<- lars(xb, yb, type = "lasso", intercept = FALSE)

plot(lasso_fit_Boys, breaks = F)

abline(v = 0.1919, col = "2")
```

#Plot 10 folds cross validation plot for boys

```
set.seed(123)

lasso_cv_Boys<- cv.lars(xb, yb, K = 10, trace = F, plot.it = T, se = T, type = "lasso")
```

#Find the optimal fraction that minimize the CV error for boys

```
op_frac_Boys<-lasso_cv_Boys$index[which.min(lasso_cv_Boys$cv)]
```

#Find the estimated coefficients of fitted LASSO regression model for boys

```
beta_Boys<- predict(lasso_fit_Boys, s = op_frac_Boys, type = "coef", mode = "fraction")$coef

beta_Boys
```

#Find the fitted values of estimated LASSO regression model for boys

```
pred_Boys<- predict(lasso_fit_Boys, xb, s = op_frac_Boys, type = "fit", mode = "fraction")$fit

pred_Boys
```

#Estimate the parameters and summary statistics of LASSO regression model for boys

```
library(lasso2)

Boyslasso<- l1ce(yb ~ xb, bound = 0.1919)

summary(Boyslasso)
```

```
#Read the girls data file, namely 'Boys_data.csv' from D
drive

GrilsData<- read.csv(file = "D:Girls_data.csv", sep = ",",
header = TRUE)

yg<- as.numeric(GrilsData [,15])

xg<- as.matrix(GrilsData[,1:14])

#Estimate the parameters of LASSO regression model for
girls

library(lars)

lasso_fit_Girls<- lars(xg, yg, type = "lasso",intercept =
FALSE)

plot(lasso_fit_Girls, breaks = F)

abline(v =0.7879, col = "2")

#Plot 10 folds cross validation plot for girls

set.seed(123)

lasso_cv_Girls<- cv.lars(xg, yg, K = 10, trace = F, plot.it
= T, se = T, type = "lasso")

#Find the optimal fraction that minimize the CV error for
girls

op_frac_Girls<- lasso_cv_Girls$index[which.min(lasso_cv_Girls$cv)]

#Find the estimated coefficients of fitted LASSO regression
model for girls

beta_Girls<- predict(lasso_fit_Girls, s = op_frac_Girls,
type = "coef", mode = "fraction")$coef
```

```
beta_Girls
```

#Find the fitted values of estimated LASSO regression model for girls

```
pred_Girls<-    predict(lasso_fit_Girls,    xg,    s    =
op_frac_Girls, type = "fit", mode = "fraction")$fit

pred_Girls
```

#Estimate the parameters and summary statistics of LASSO regression model for girls

```
library(lasso2)

Girlslasso<- l1ce(yg ~ xg, bound = 0.7879)

summary(Girlslasso)
```

CHAPTER FIVE

RESULTS AND DISCUSSION

"If Data Analysis is to be well Done Mach of it must be a Matter of Judgment, and "Theory" whether Statistical or Non Statistical Well Have to Guide Not Command." - J.W. Tukey Chilli's

CHAPTER 5

RESULTS AND DISCUSSION

5.1 Outline

This chapter includes descriptive statistics for the model parameters and that of curve fitting, performance of the BTT and that of the proposed model, velocity and acceleration curves, weight estimation at final stature, predicted final stature and predicting equations for final stature with their precession, validity and stability.

5.2 Descriptive Statistics of Model Parameters

The STATISTICA 8.0 software was used to estimate the parameters for both BTT model (Eq. 4.1) and proposed model (Eq. 4.3) separately to the 300 sets of longitudinal data (described in materials section) using Bayesian method (described in methods section) and accumulated the parameters shown Table 5.1a and Table 5.1b. Here, the maximum number of iterations and convergence criterion were used, respectively, as 10000 and 0.0000009999. For BTT model, on average 34.097%, 4.757%, and 61.146% of the total final stature were completed during early, middle and adolescent phase of growth, respectively, for the male population and for the female population, these percentages were 21.929%, 23.829%, and 54.242%, respectively (Table 5.1a). For proposed model, on average 30.102%, 29.933%, and 39.965% of the total final stature were completed during early, middle and adolescent phase of growth, respectively, for the male population and for the female population, these percentages were 29.169%, 36.137%, and 34.694%, respectively (Table 5.1b). The correlation coefficient between different parameters of the BTT and proposed model were shown in the same Table 5.1a and Tale 5.1b. For BTT model (Table 5.1a), the correlation coefficient between different parameters (a_1, b_1), (a_1, c_1), (a_1, a_3), (b_1, c_1), (a_2, a_3), (b_2, c_2) and (b_3, c_3) for boys, and that of between (a_1, b_1), (a_1, c_1), (a_1, a_2), (b_1, c_1), (a_2, b_2), (a_2, c_2), (b_2, c_2) and (b_3, c_3) for girls were statistically significant ($p \leq 0.05$). This table also implied that the average root mean square error of the estimate for boys was larger than that for girls. Similarly, for the proposed model (Table 5.1b), the correlation coefficient between different parameters (a_1, a_2), (a_1, a_{21}), (a_1, a_{22}), (a_1, c_2), (a_{11}, a_{12}), (a_{11}, c_1), (a_{12}, c_1), (a_{21}, a_{22}), (a_{21}, c_2), (a_{22}, c_2), (a_{31}, a_{32}), (a_{31}, c_3) and (a_{32}, c_3) for boys, and that of between (a_1, a_2), (a_1, a_3), (a_{11}, a_{12}), (a_{11}, c_1), (a_{12}, c_1), (a_{12}, a_3), (a_3, c_1), (a_2, a_{22}), (a_2, c_2), (a_{21}, c_2), (a_3, c_3), (a_{31}, a_{32}) and (a_{32}, c_3) for girls were statistically significant ($p \leq 0.05$). Like Table 5.1a, average root mean square error of the estimate for boys was larger than that for girls here also. Average root mean square error of the proposed model was smaller than that of the BTT model for both boys and girls (Table 5.1a and Table 5.1b).

Table 5.1a Estimated population mean, standard deviation (SD), and correlation matrix of BTT model parameters for boys and girls

Parameter		a_1	b_1	c_1	a_2	b_2	c_2	a_3	b_3	c_3
					Boys					
Mean (N= 180)		48.743	3.718	-29.675	6.801	5.124	-40.049	87.410	3.366	9.791
SD		7.760	1.275	0.039	6.404	0.680	1.928	8.299	2.776	0.929
Correlation	a_1									
Matrix	b_1	-0.212(**)								
	c_1	0.193(**)	-0.901(**)							
	a_2	0.086	0.027	-0.030						
	b_2	0.017	-0.069	0.057	0.002					
	c_2	-0.028	0.073	-0.062	-0.001	-0.910(**)				
	a_3	-0.527(**)	0.087	-0.085	-0.241(**)	0.128	-0.124			
	b_3	0.063	-0.039	0.027	0.004	-0.036	0.035	-0.112		
	c_3	0.135	0.001	0.011	0.000	-0.014	0.011	-0.071	-0.763(**)	
Average root mean square error of the estimate: 0.814738										
					Girls					
Mean (N = 120)		43.449	4.813	-23.124	47.213	4.620	-27.469	107.471	7.938	-6.728
SD		6.689	2.338	0.418	5.798	0.561	1.666	7.998	2.099	1.210
Correlation	a_1									
Matrix	b_1	-0.212(*)								
	c_1	0.312(**)	-0.927(**)							
	a_2	-0.302(**)	0.031	-0.027						
	b_2	0.043	-0.078	0.057	-0.216(*)					
	c_2	-0.124	0.074	-0.061	0.241(**)	-0.952(**)				
	a_3	-0.115	-0.002	0.003	-0.038	-0.009	0.013			
	b_3	0.151	0.000	0.008	0.028	-0.134	0.090	-0.019		
	c_3	-0.123	0.002	-0.007	0.019	0.094	-0.051	-0.001	-0.886(**)	
Average root mean square error of the estimate: 0.760471										

Note: The symbols *and ** indicate that the correlation are statistically significant at the 0.05 (p = 0.05) and 0.01 (p = 0.01) level (2-tailed), respectively.

Table 5.1b Estimated population mean, standard deviation (SD), and correlation matrix of proposed model parameters for boys and girls

Parameter		a_1	a_{11}	a_{12}	c_1	a_1	a_{21}	a_{22}	c_2	a_3	a_{31}	a_{32}	c_3
							Boys						
Mean (N=180)		57.352	-26.367	14.330	-98.300	57.032	2.639	3.105	-61.723	76.146	7.014	-0.827	-4.926
SD		6.551	1.632	1.314	0.289	5.761	4.023	2.097	3.058	9.903	2.920	1.748	0.602
Correlation Matrix	a_1												
	a_{11}	0.119											
	a_{12}	-0.122	-0.982(**)										
	c_1	0.135	0.998(**)	-0.998(**)									
	a_2	-0.567(**)	0.055	-0.053	0.041								
	a_{21}	-0.155(*)	0.003	-0.002	0.000	-0.018							
	a_{22}	0.169(*)	0.006	-0.006	0.008	-0.040	-0.887(**)						
	c_2	-0.154(*)	-0.011	0.012	-0.014	0.110	0.680(**)	-0.933(**)					
	a_3	-0.030	-0.037	0.036	-0.035	-0.102	0.030	-0.017	-0.002				
	a_{31}	-0.069	0.008	-0.007	0.003	0.138	-0.001	-0.003	0.007	-0.022			
	a_{32}	0.080	-0.004	0.003	0.002	-0.140	0.000	0.001	-0.001	0.015	-0.992(**)		
	c_3	-0.084	0.000	0.001	-0.006	0.117	0.000	0.002	-0.005	-0.005	0.970(**)	-0.986(**)	
Average root mean square error of the estimate: 0.653058													
							Girls						
Mean (N=120)		48.479	2.333	1.153	-27.639	60.060	6.375	-1.910	12.511	57.662	11.101	1.625	-39.635
SD		6.151	2.026	1.639	0.381	5.902	4.951	1.894	2.864	4.941	1.527	2.787	1.113
Correlation Matrix	a_1												
	a_{11}	-0.050											
	a_{12}	-0.032	-0.918(**)										
	c_1	0.144	0.657(**)	-0.888(**)									
	a_2	-0.511(**)	0.041	-0.025	0.027								
	a_{21}	0.052	-0.007	-0.016	0.040	-0.169							
	a_{22}	-0.012	0.008	0.008	-0.025	0.138	-0.961(**)						
	c_2	-0.020	-0.010	-0.002	0.014	-0.094	0.891(**)	-0.976(**)					
	a_3	-0.404(**)	-0.105	0.183(*)	-0.285(**)	-0.454(**)	0.079	-0.108	0.109				
	a_{31}	0.064	-0.036	0.023	-0.015	-0.116	0.004	-0.001	-0.006	0.044			
	a_{32}	-0.046	0.059	-0.057	0.058	0.125	-0.026	0.017	-0.003	-0.086	-0.682(**)		
	c_3	-0.098	-0.064	0.087	-0.115	-0.071	0.041	-0.033	0.017	0.201(*)	0.092	-0.716(**)	
Average root mean square error of the estimate: 0.608818													

Note: The symbols * and ** indicate that the correlation are statistically significant at the 0.05 (p = 0.05) and 0.01 (p = 0.01) level (2-tailed), respectively.

5.3 Average Curve Fitting

The average structural curve fitted values of stature for boys and girls with their stature differences, obtained from fitted BTT and proposed model, for every year from age 1 to 25 years were presented in Table 5.2. This results (Table 5.2) showed that the distribution of predicted stature, on average, the boys became taller than girls from age 1 to 9 and 12 to 25. But, the distribution of predicted stature, on average, showed that the girls became taller than boys from age 10 to 11.

Table 5.2 The average structural curve fitted values of stature for boys and girls with their stature differences (boys less girls) by age (year) using BTT and proposed model

Age (year)	Stature (cm) of Boys		Stature (cm) of Girls		Difference in Stature	
	BTT	**Proposed**	**BTT**	**Proposed**	**BTT**	**Proposed**
1	73.694	74.260	73.532	73.618	0.162	0.642
2	85.268	85.024	84.331	84.337	0.937	0.687
3	93.865	94.073	92.806	92.823	1.059	1.250
4	100.920	101.052	100.278	100.291	0.642	0.761
5	107.324	107.162	106.795	106.722	0.529	0.440
6	113.700	113.612	112.862	112.863	0.838	0.749
7	119.813	119.819	118.947	118.908	0.866	0.911
8	125.696	125.757	124.604	124.695	1.092	1.062
9	131.127	131.203	130.314	130.351	0.813	0.852
10	136.424	136.425	136.461	136.446	-0.037	-0.021
11	142.009	141.957	142.796	142.642	-0.787	-0.685
12	148.463	148.448	148.297	148.418	0.166	0.030
13	156.020	155.994	152.383	152.437	3.637	3.557
14	162.669	160.405	155.088	155.144	7.581	5.261
15	167.259	165.438	156.580	156.550	10.679	8.888
16	169.624	168.752	157.423	157.366	12.201	11.386
17	170.837	170.963	157.893	157.841	12.944	13.122
18	171.592	171.102	158.113	158.171	13.479	12.931
19	171.910	171.330	158.117	158.200	13.793	13.130
20	171.923	171.412	158.388	158.320	13.535	13.092
21	171.932	171.476	158.490	158.335	13.442	13.141
22	171.941	171.481	158.500	158.346	13.441	13.135
23	171.946	171.494	158.521	158.347	13.425	13.147
24	171.951	171.503	158.532	158.348	13.419	13.155
25	171.952	171.513	158.533	158.349	13.419	13.164

5.4 Comparisons Between the BTT Model and the Proposed Model

The performance of the BTT and proposed model could be visually confirmed by plotting observed and fitted statures that shown in Figures 5.1a and 5.1b (for a single data set, say data set number 1) for boys and girls, respectively. Here, the shape parameters were assumed both models as d_1= 0.75, d_2= 0.75 and d_3= 1.20. Similar graph can be shown for all other data sets for respective boys and girls but this line graphs were omitted because the performance of these models could easily be compared by mean square error. Thus, the mean square errors of the BTT and proposed model were computed for all data set. A single data set (data set number 1) was used to compare the precession of the BTT and proposed model due to changes of the shape parameters (Figure 5.2a for boys and Figure 5.2b for girls). These figures always divulge the superiority of the proposed model. On the other hand, the precessions of the BTT and proposed model was tested for all the 300 data sets shown in Figure 5.3a (Shown also in Table 1 within the Appendix-2) for boys and Figure 5.3b (Shown also in Table 2 within the Appendix-2) for girls and found that the proposed model was more precise everywhere.

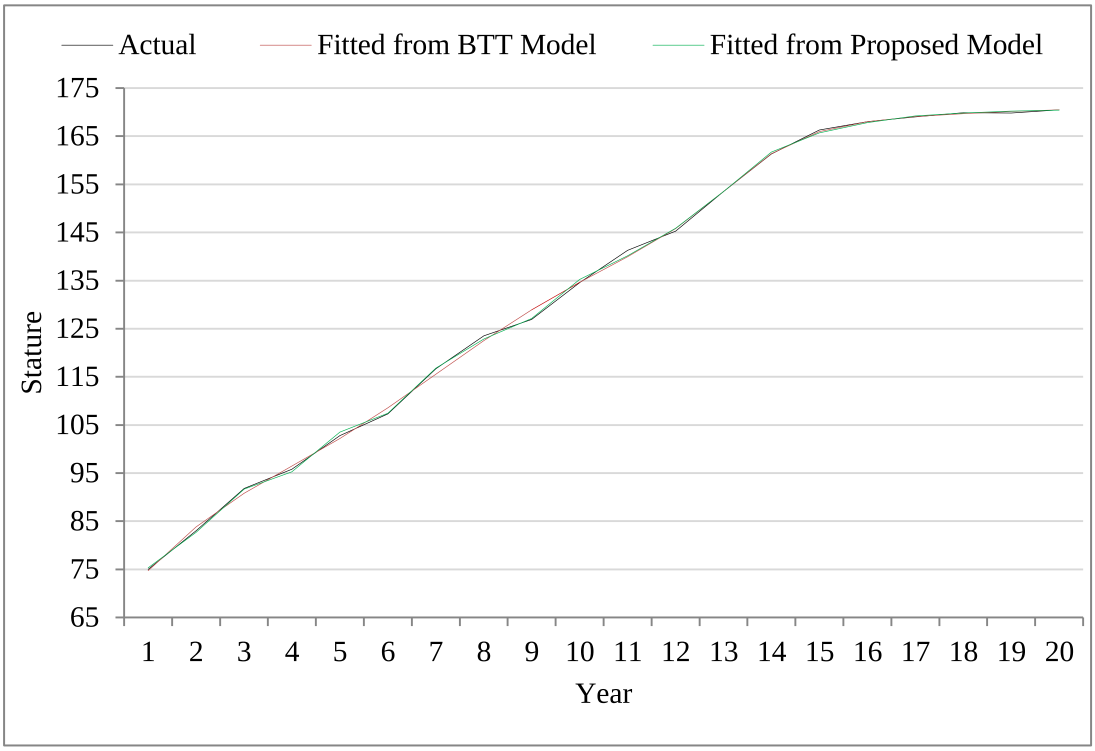


Figure 5.1a Line graph for actual stature from the observed; fitted stature from BTT and proposed models of boys

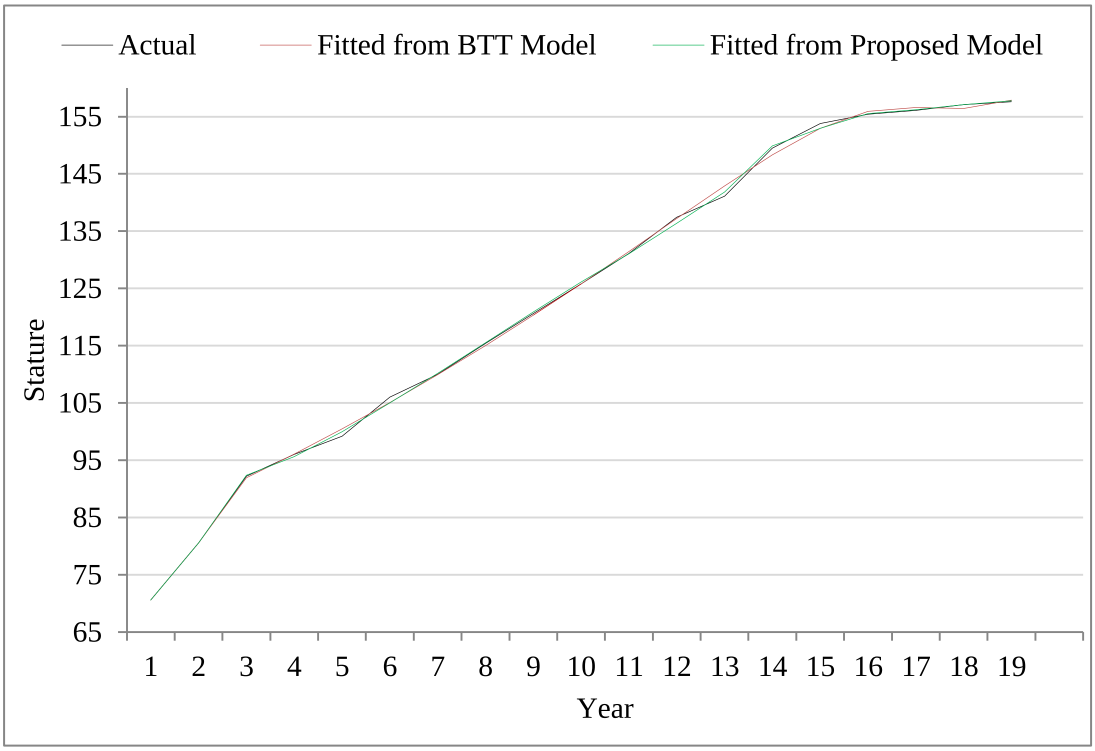


Figure 5.1b Line graph for actual stature from the observed; fitted stature from BTT and proposed models of girls

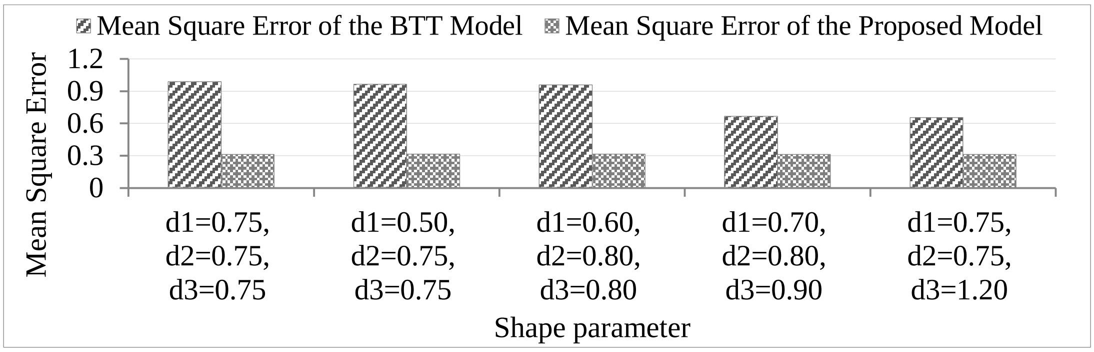


Figure 5.2a Clustered column bar diagram for the mean square error of BTT and proposed model of boys for data set number 1

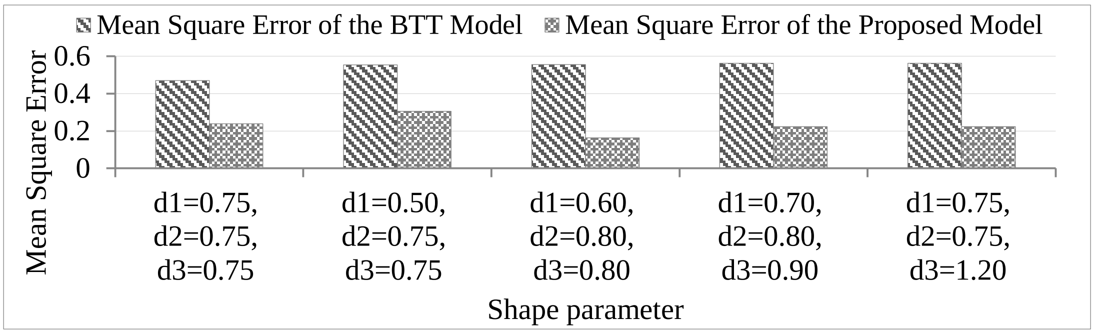
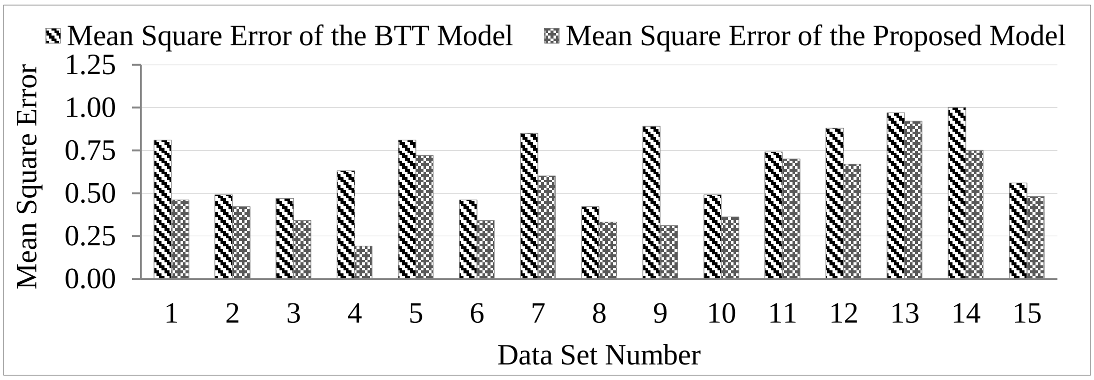
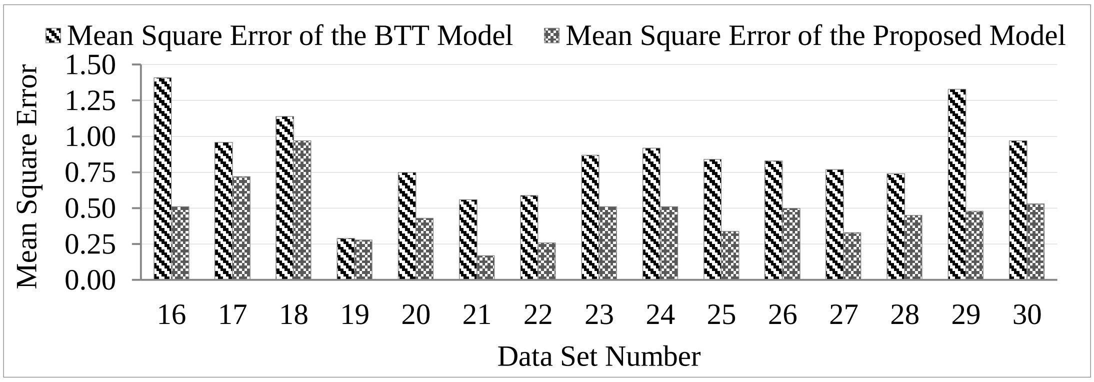


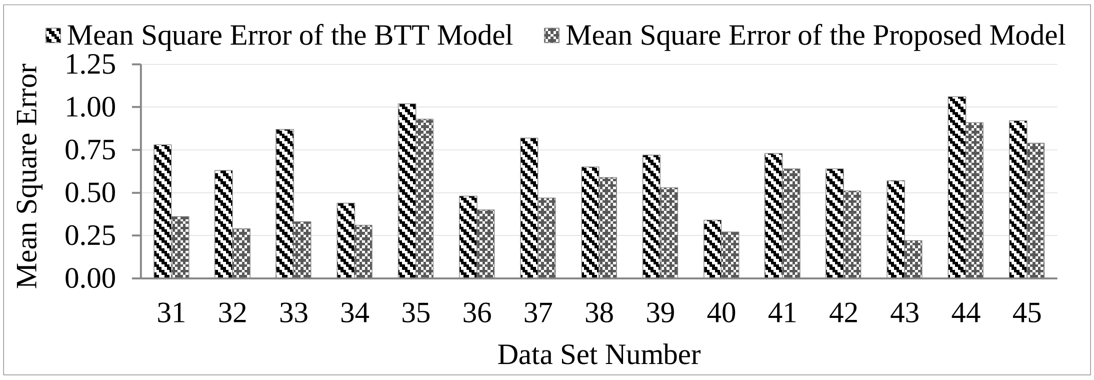
Figure 5.2b Clustered column bar diagram for the mean square error of BTT and proposed model of girls for data set number 1



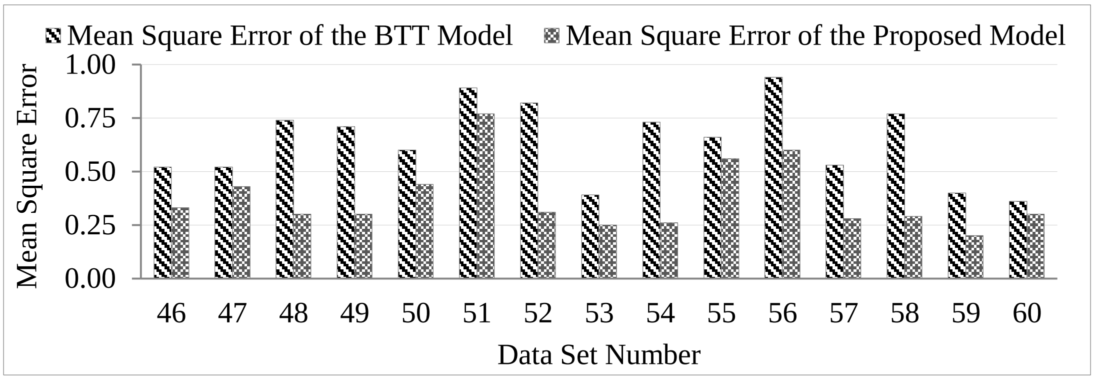
(1)



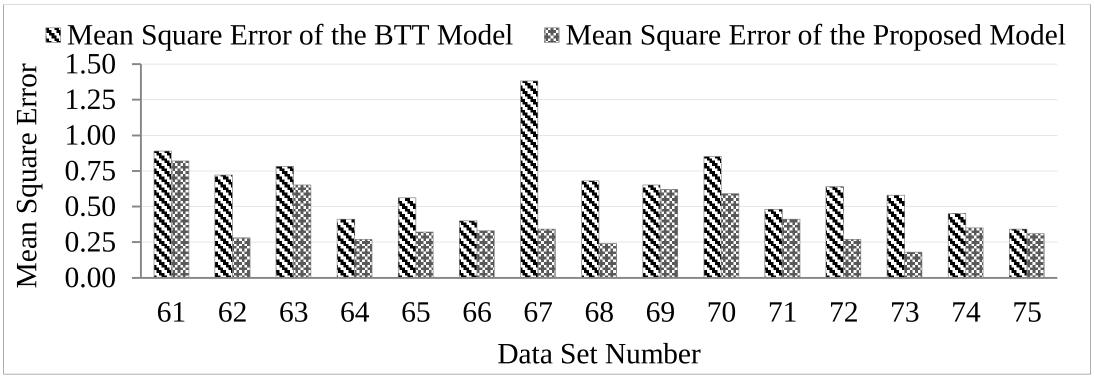
(2)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.25
1.00
0.75
0.50
0.25
0.00
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
Data Set Number

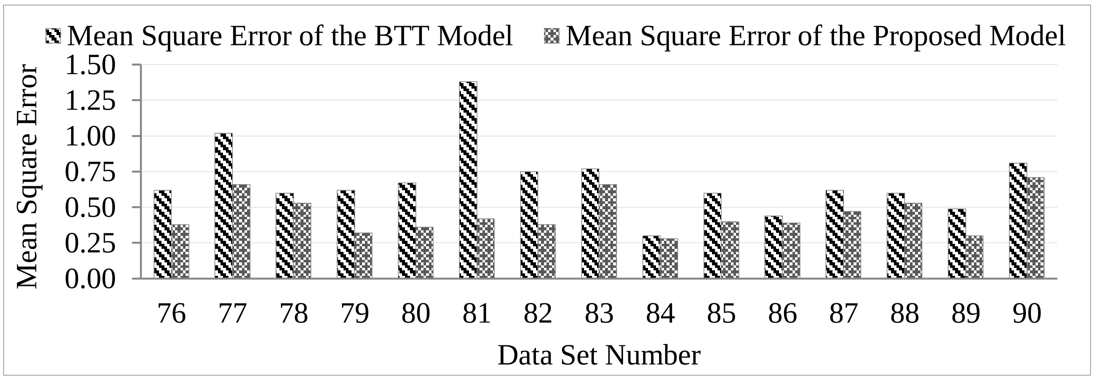
(3)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.00
0.75
0.50
0.25
0.00
46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
Data Set Number

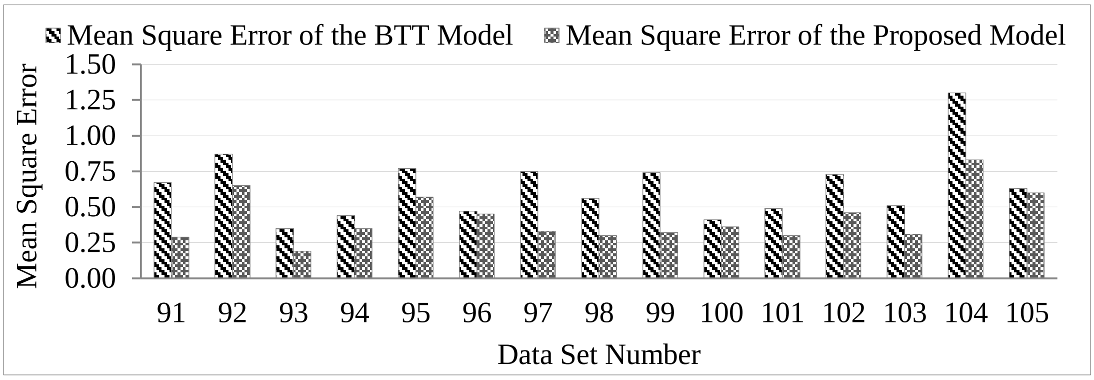
(4)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.50
1.25
1.00
0.75
0.50
0.25
0.00
61 62 63 64 65 66 67 68 69 70 71 72 73 74 75
Data Set Number

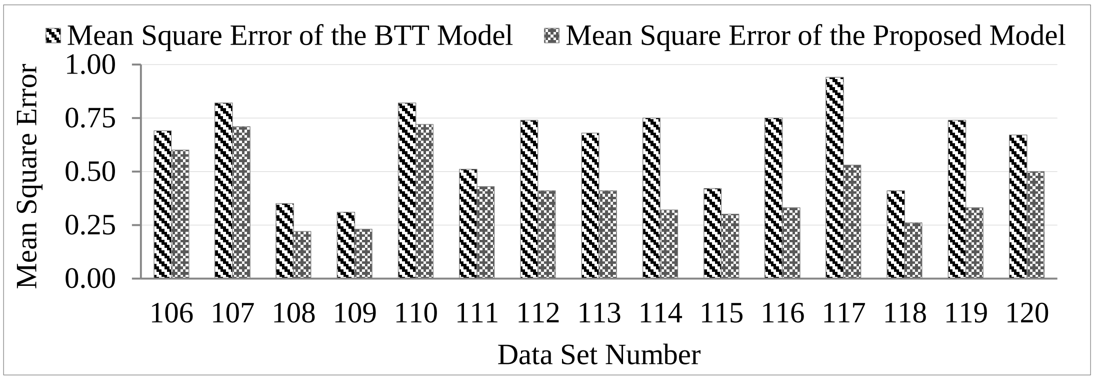
(5)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.50
1.25
1.00
0.75
0.50
0.25
0.00
76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
Data Set Number

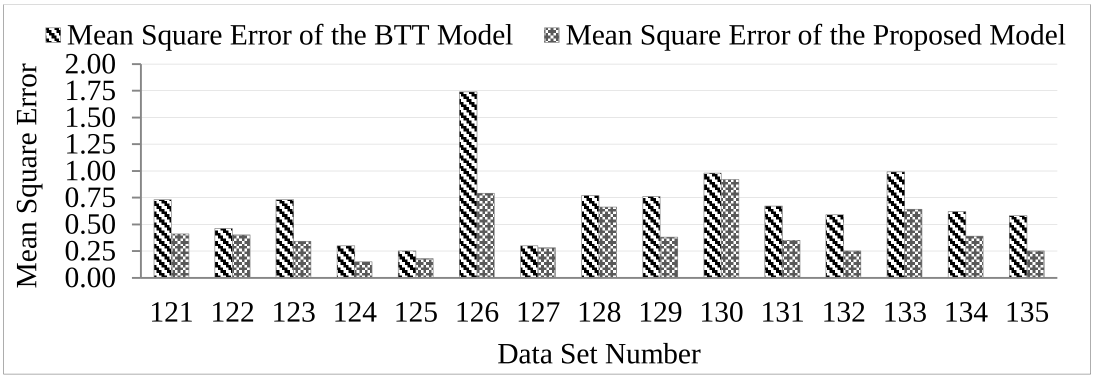
(6)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.50
1.25
1.00
0.75
0.50
0.25
0.00
91 92 93 94 95 96 97 98 99 100 101 102 103 104 105
Data Set Number

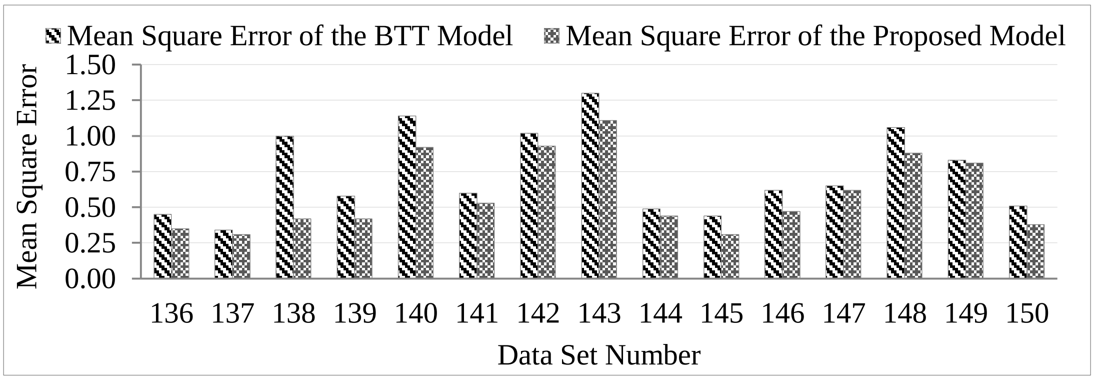
(7)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.00
0.75
0.50
0.25
0.00
106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
Data Set Number

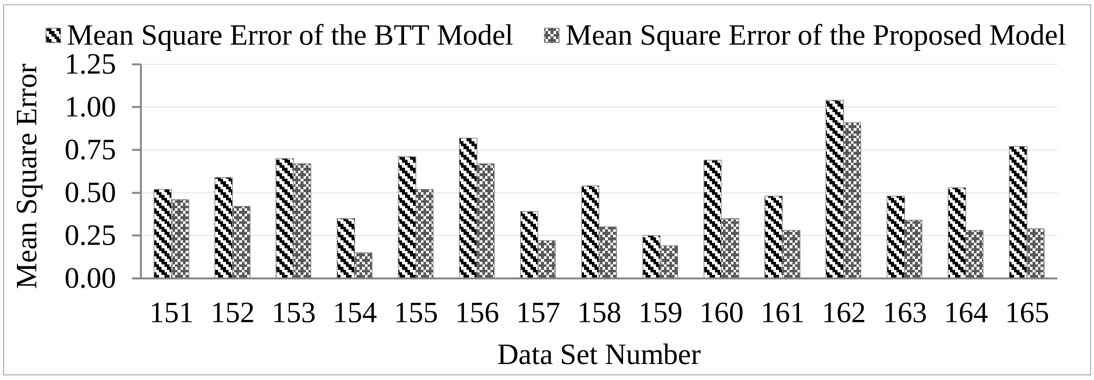
(8)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
121 122 123 124 125 126 127 128 129 130 131 132 133 134 135
Data Set Number

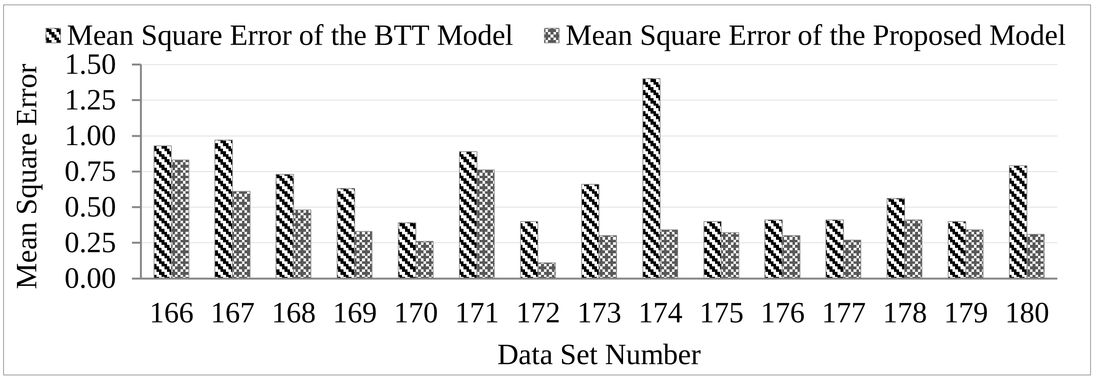
(9)


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.50
1.25
1.00
0.75
0.50
0.25
0.00
136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
Data Set Number

(10)

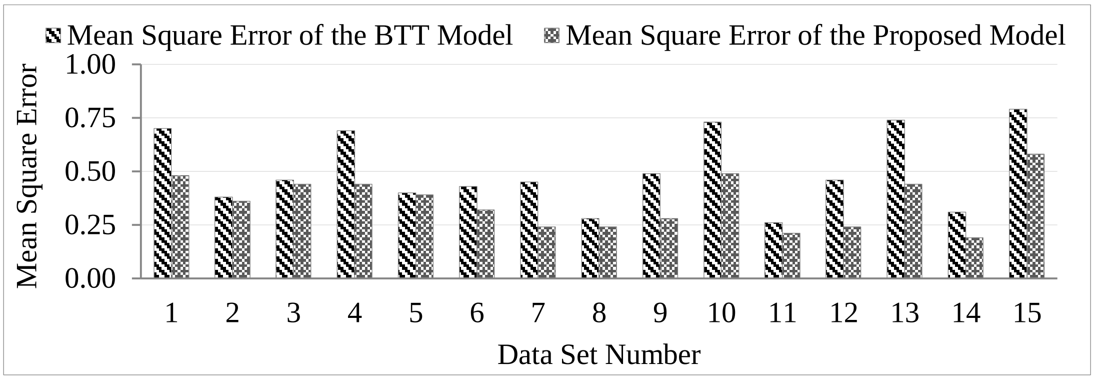


(11)



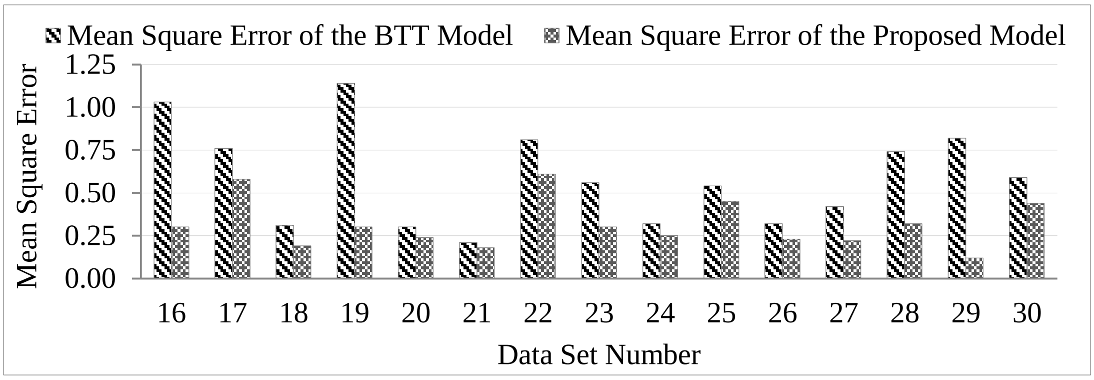
(12)

Figure 5.3a Clustered column bar diagram (1−12) for the mean square error of BTT and proposed models of boys



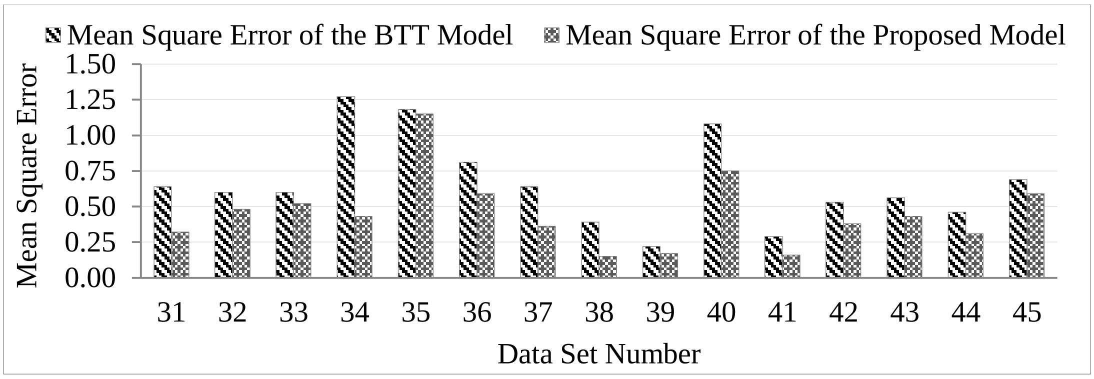
Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.00
0.75
0.50
0.25
0.00
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
Data Set Number

(1)



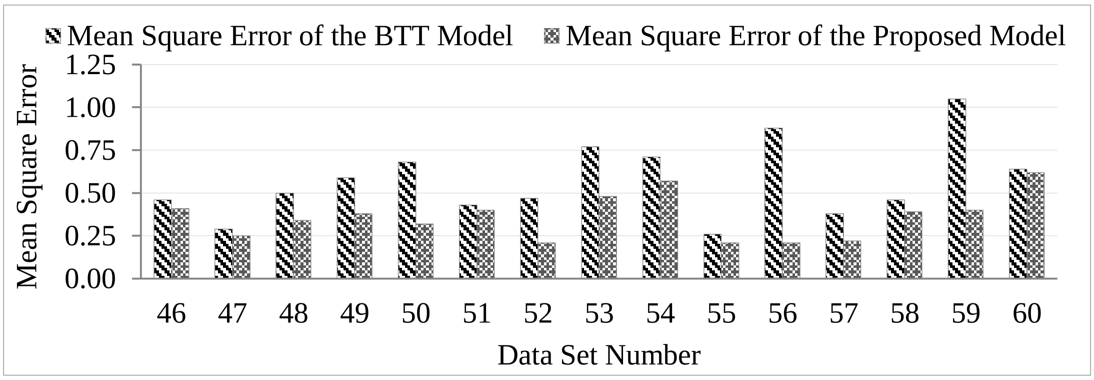
Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.25
1.00
0.75
0.50
0.25
0.00
16 17 18 19 20 21 22 23 24 25 26 27 28 29 30
Data Set Number

(2)



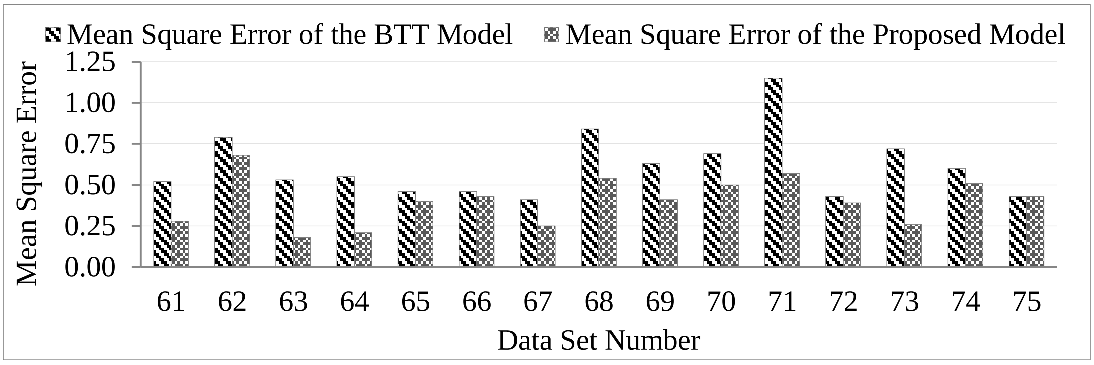
Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.50
1.25
1.00
0.75
0.50
0.25
0.00
31 32 33 34 35 36 37 38 39 40 41 42 43 44 45
Data Set Number

(3)

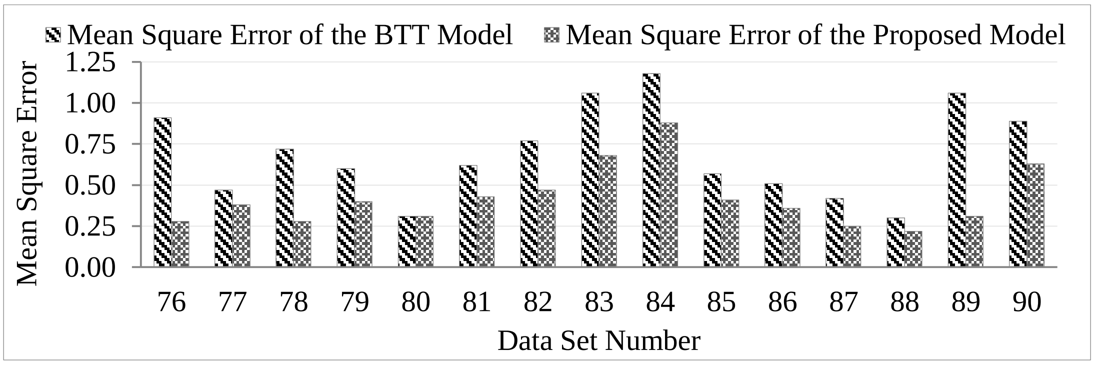


Mean Square Error of the BTT Model
Mean Square Error of the Proposed Model
Mean Square Error
1.25
1.00
0.75
0.50
0.25
0.00
46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
Data Set Number

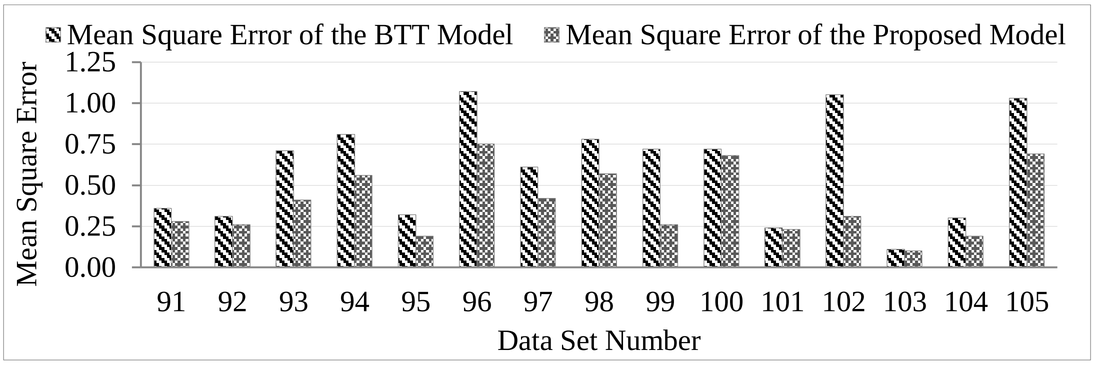
(4)



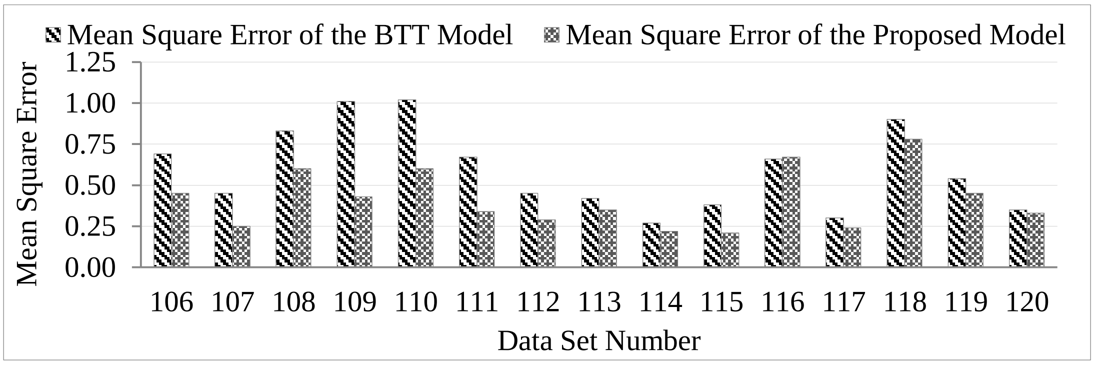
(5)



(6)



(7)



(8)

Figure 5.3b Clustered column bar diagram (1–8) for the mean square error of BTT and proposed models of girls

5.5 Velocity, Acceleration, Minimum and Maximum Points of the Proposed Model

For data set number 1, for example, velocity, acceleration, local maxima, local minima and saddle points were calculated accordingly (details are methods section) and shown in Tables 5.3a and 5.3b for boys and girls, respectively. Table 5.3a showed one local minimum point, twelve local maximum points and seven saddle points for boys. Table 5.3b showed four local minimum points, seven local maximum points and seven saddle points for girls. These local minima and local maxima can be used as biological parameters.

Table 5.3a Velocity and acceleration of proposed model for boys

Stature	Predicted	F*t*	F*x*	F*tt*	F*xx*	F*tx*	Det(H)	Comment
75.00	75.32	-10.395	10.516	20.645	4.634	-10.067	-5.681	Saddle
83.00	82.70	-6.400	8.041	25.820	5.793	-12.585	-8.796	Saddle
91.80	91.66	-5.407	6.805	21.812	4.433	-10.275	-8.890	Saddle
95.80	95.29	0.484	4.347	-36.623	-11.832	20.600	8.969	L. Max
102.80	103.49	3.176	2.355	-36.454	-11.962	20.670	8.815	L. Max
107.30	107.42	-7.777	8.333	-23.953	-8.473	14.102	4.090	L. Max
116.70	116.85	8.895	-1.582	-34.502	-11.269	19.640	3.065	L. Max
123.50	122.88	3.297	1.530	-35.088	-11.393	20.012	-0.706	Saddle
126.90	127.12	17.190	-6.094	-15.378	-5.474	9.405	-4.275	Saddle
134.60	135.27	13.166	-3.348	-36.565	-11.385	20.695	-12.000	Saddle
141.30	140.17	21.962	-7.542	-11.624	-4.068	7.222	-4.865	Saddle
145.30	145.86	23.031	-7.716	19.436	4.995	-9.751	2.005	L. Min
153.50	153.50	5.668	0.766	-0.932	-0.041	-0.078	0.032	L. Max
161.30	161.69	3.785	0.477	-1.241	-0.022	-0.150	0.005	L. Max
166.30	165.67	2.322	0.292	-0.964	-0.016	-0.117	0.002	L. Max
168.00	167.83	1.377	0.174	-0.640	-0.011	-0.077	0.001	L. Max
169.00	169.23	0.713	0.090	-0.354	-0.006	-0.043	0.000	L. Max
169.90	169.84	0.405	0.052	-0.208	-0.004	-0.025	0.000	L. Max
169.80	170.22	0.213	0.028	-0.111	-0.002	-0.013	0.000	L. Max
170.50	170.43	0.105	0.013	-0.055	-0.001	-0.007	0.000	L. Max

Note: F*t*, F*x*; F*tt*, F*xx*; F*tx* and Det(H) denote the velocity with respect to age and weight; acceleration with respect to age and weight; the partial derivative with respect to age and weight and determinant of 2×2 Hessian matrix, respectively. All L. Max are equal to local maximum in table in above.

Table 5.3b Velocity and acceleration of proposed model for girls

Stature	Predicted	F*t*	F*x*	F*tt*	F*xx*	F*tx*	Det(H)	Comment
70.60	70.60	-244.804	171.046	-2532.180	-1220.007	1757.812	-0.148	Saddle
80.60	80.60	-148.291	104.466	-2025.334	-975.997	1406.187	0.000	Saddle
92.20	92.36	3.034	-0.242	1.340	0.114	-0.289	0.000	L. Min
96.00	95.62	4.185	-0.435	1.561	0.126	-0.333	0.000	L. Min
99.20	99.95	5.447	-0.653	1.473	0.115	-0.304	0.000	L. Min
106.00	105.03	6.369	-0.770	0.504	-0.123	0.115	0.000	Saddle
110.00	110.16	6.734	-0.746	0.267	0.105	-0.096	0.000	L. Min
115.40	115.51	6.544	-0.570	-0.558	0.093	0.053	0.000	Saddle
120.60	120.91	5.838	-0.284	-1.278	0.080	0.185	0.000	Saddle
125.80	126.13	4.963	0.147	-1.611	0.076	0.245	0.000	Saddle
131.10	131.04	3.708	0.530	-1.734	0.064	0.272	0.000	Saddle
137.40	136.35	2.780	0.930	-1.645	0.018	0.279	0.000	Saddle
141.10	141.82	2.037	1.149	-1.489	-0.069	0.293	0.000	L. Max
149.50	149.89	2.928	0.329	-1.669	-0.213	0.394	0.000	L. Max
153.80	152.96	2.058	0.329	-1.364	-0.189	0.329	0.000	L. Max
155.40	155.52	1.756	0.099	-1.158	-0.141	0.270	0.000	L. Max
156.10	156.19	0.573	0.524	-0.624	-0.151	0.181	0.000	L. Max
157.10	157.11	0.226	0.538	-0.398	-0.136	0.135	0.000	L. Max
157.60	157.76	0.024	0.527	-0.252	-0.123	0.103	0.000	L. Max

Note: F*t*, F*x*; F*tt*, F*xx*; F*tx* and Det(H) denote the velocity with respect to age and weight; acceleration with respect to age and weight; the partial derivative with respect to age and weight and determinant of 2×2 Hessian matrix, respectively. All L. Max are equal to local maximum in table in above.

5.6 Predictions of Final Stature

According to the Bock et al. (1994), stature at the age 25 year had been considered as a predicted final stature (*PFS*). However, the definition of age at final stature was different by researchers (Kato et al., 1998). The age was the only independent variable in BTT model and the value of age at final stature was known exactly. But, the proposed model had two independent variables they were (i) age and (ii) weight. Here, weight at age 25 needs to be estimated. For this, the following steps may be considered to predict the final stature.

5.6.1 First Stage

In this stage, the weight at final stature was estimated. The statistical software R i386 3.0.1 with the package '*forecast'* was used to choose the optimum order of ARIMA model as well as the optimum value of Box–Cox parameter (λ) and the remaining models were estimated by the STATISTICA 8.0 software. The summary of ARIMA model, values of Box-Cox parameter (λ), MSE of ARIMA with λ, MSE of Gompertz, MSE of logistic, MSE of double logistic models and predicted weight for boys and girls were presented in Tables 5.4a and 5.4b, respectively, where the selected best model indicated by asterisk. The estimated weight at the final stature of each data set for both boys and girls were presented in the last column of Tables 5.4a and 5.4b, respectively.

Table 5.4a Weight estimation using ARIMA model with Box-Cox transformation, Gompertz, logistic and double logistic models for boys

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error				Predicted Weight
			ARIMA with λ	Gomp–ertz	Logistic	Double logistic	
01	ARIMA(0,1,0) with drift	0.23	2.34	2.00	1.32	0.36*	63.06
02	ARIMA(0,1,0) with drift	0.13	2.33	1.93	1.76	0.96*	60.34
03	ARIMA(1,1,0) with drift	0.70	2.33	1.99	1.56	1.08*	57.85
04	ARIMA(0,1,0) with drift	1.11	3.77	3.17	2.70	2.01*	76.19
05	ARIMA(0,1,0) with drift	0.74	1.20	1.05	0.92	0.65*	64.52
06	ARIMA(1,1,0) with drift	1.51	2.56	3.16	3.03	1.88*	77.42
07	ARIMA(0,1,0) with drift	0.20	1.67	1.55	1.35*	2.16	58.99
08	ARIMA(0,2,0)	0.35	2.06	2.26	1.83	0.50*	66.86
09	ARIMA(0,1,0) with drift	0.24	1.85	1.83	1.67	0.56*	55.65
10	ARIMA(0,1,0) with drift	0.08	1.70	1.39	1.28	0.63*	65.13
11	ARIMA(0,1,0) with drift	0.86	2.30	1.58	1.70	1.20*	64.04
12	ARIMA(0,1,0) with drift	0.99	2.12	2.29	2.29	1.23*	61.76
13	ARIMA(0,1,0) with drift	0.73	1.34	1.35	1.25	0.50*	65.44
14	ARIMA(1,1,0) with drift	0.58	2.23	2.85	2.38	1.12*	73.24
15	ARIMA(0,1,0) with drift	0.70	0.78	0.72	0.69	0.44*	73.50
16	ARIMA(0,2,1)	0.09	2.74	1.40	1.34	1.23*	69.39
17	ARIMA(0,1,0) with drift	0.62	1.95	1.56	1.36	0.88*	59.84
18	ARIMA(1,1,0) with drift	0.46	3.39	2.72	2.71	1.88*	62.95
19	ARIMA(0,2,0)	-0.26	1.90	2.14	1.83	0.50*	51.82

Continued…

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error ARIMA with λ	Gomp–ertz	Logistic	Double logistic	Predicted Weight
20	ARIMA(0,1,1) with drift	0.06	1.32	1.72	1.50	0.22*	56.83
21	ARIMA(1,1,0) with drift	0.53	2.16	2.65	2.06	0.82*	65.69
22	ARIMA(0,2,0)	-0.14	3.05	2.72	2.35	0.79*	57.50
23	ARIMA(0,1,0) with drift	0.08	4.12	2.42	2.28	1.99*	55.68
24	ARIMA(0,1,0) with drift	0.83	1.77	1.87	1.77	0.56*	59.82
25	ARIMA(0,1,1) with drift	0.59	1.49	1.99	1.68	1.02*	64.75
26	ARIMA(0,1,0) with drift	0.32	2.43	2.49	2.35	0.71*	60.48
27	ARIMA(0,1,0) with drift	0.22	1.60	1.48	1.30	0.48*	52.44
28	ARIMA(0,1,0) with drift	0.56	2.06	2.42	2.42	0.67*	57.03
29	ARIMA(1,1,0) with drift	0.28	1.80	2.17	1.85	0.58*	59.66
30	ARIMA(0,1,0) with drift	0.56	1.44	1.33	1.33	0.76*	65.09
31	ARIMA(0,2,0)	-0.10	2.65	2.34	1.93	0.84*	65.23
32	ARIMA(0,2,0)	0.14	1.83	1.84	1.53	0.67*	62.06
33	ARIMA(1,2,0)	-0.06	2.34	2.30	2.04	0.57*	69.44
34	ARIMA(0,1,0) with drift	0.77	1.23	0.99	0.93	0.78*	65.81
35	ARIMA(1,1,0) with drift	0.65	1.42	2.09	1.89	0.63*	66.49
36	ARIMA(1,1,0) with drift	0.19	2.21	2.75	2.33	0.89*	63.86
37	ARIMA(0,2,0)	0.39	1.84	2.68	2.13	1.37*	63.90
38	ARIMA(0,2,1)	0.29	6.21	5.27	4.77	4.28*	81.28
39	ARIMA(1,1,0) with drift	0.27	1.55	1.74	1.34	0.57*	67.10
40	ARIMA(1,1,0) with drift	0.50	1.19	1.83	1.51	0.45*	58.69
41	ARIMA(0,1,0) with drift	0.57	1.08	0.87	0.71	0.42*	59.72
42	ARIMA(1,2,0)	-0.11	2.06	1.31	1.30	0.61*	63.04
43	ARIMA(1,2,0)	-0.30	3.41	1.49	1.31	1.21*	62.80
44	ARIMA(0,2,0)	-0.08	2.97	2.84	2.38	0.72*	59.77
45	ARIMA(0,2,0)	0.53	1.96	1.68	1.37	1.23*	65.37
46	ARIMA(0,1,0) with drift	0.23	2.14	1.79	1.53	0.70*	65.50
47	ARIMA(1,2,0)	-0.19	4.00	2.28	2.15	2.05*	73.66
48	ARIMA(0,1,0) with drift	0.93	1.89	1.39	1.41	1.09*	78.87
49	ARIMA(0,1,0) with drift	0.30	1.93	2.05	1.89	0.70*	55.92
50	ARIMA(1,1,0) with drift	0.32	2.00	2.55	2.33	0.62*	61.93
51	ARIMA(0,1,0)	2.90	16.26	9.85	10.30	8.06*	69.50
52	ARIMA(2,1,0) with drift	0.54	2.19	3.82	3.24	1.98*	56.11
53	ARIMA(1,1,0) with drift	0.07	1.73	2.27	1.79	0.44*	73.00
54	ARIMA(0,1,0) with drift	1.10	1.75	1.31	1.30	0.63*	90.98
55	ARIMA(0,1,0) with drift	0.44	1.84	1.48	1.32	0.95*	78.89
56	ARIMA(0,1,0) with drift	-0.02	2.80	1.94	1.67	0.93*	92.78
57	ARIMA(0,2,0)	0.28	1.98	2.52	2.04	0.72*	68.70
58	ARIMA(0,1,0) with drift	0.49	2.12	1.73	1.59	1.00*	63.24
59	ARIMA(0,1,0) with drift	0.60	3.51	2.60	2.28	2.13*	68.78
60	ARIMA(0,1,0) with drift	0.71	2.20	2.12	1.76	0.97*	59.12
61	ARIMA(0,1,0) with drift	0.21	4.40	4.21	3.53	1.53*	70.93
62	ARIMA(0,1,0) with drift	0.73	1.65	1.65	1.55	0.82*	79.85
63	ARIMA(0,1,1) with drift	-0.08	1.82	2.23	1.97	0.98*	84.97
64	ARIMA(0,2,1)	-0.04	2.43	1.71	1.35	0.83*	68.45
65	ARIMA(0,1,0) with drift	0.34	2.18	1.92	1.55	0.81*	111.34
66	ARIMA(0,1,0) with drift	0.31	1.59	1.54	1.53	0.88*	68.65
67	ARIMA(0,1,0) with drift	0.62	1.84	1.90	1.56	0.83*	58.81

Continued…

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error ARIMA with λ	Gomp–ertz	Logistic	Double logistic	Predicted Weight
68	ARIMA(0,1,0) with drift	0.92	2.93	2.27	1.91	1.39*	76.93
69	ARIMA(0,1,0) with drift	0.42	3.06	1.95*	1.95*	2.22	141.81
70	ARIMA(1,1,0)	0.44	1.90	2.30	2.05	1.55*	66.25
71	ARIMA(0,1,0) with drift	0.33	1.24	1.31	1.12	0.43*	56.22
72	ARIMA(0,1,0) with drift	0.32	2.47	2.35	2.11	0.92*	54.32
73	ARIMA(0,1,0) with drift	0.61	1.91	1.98	1.54	0.71*	62.06
74	ARIMA(0,1,0) with drift	0.69	2.51	1.60	1.29	1.18*	72.39
75	ARIMA(1,2,0)	0.21	2.08	1.88	1.28	0.84*	62.90
76	ARIMA(1,2,0)	0.31	3.41	3.95	3.50	3.06*	60.63
77	ARIMA(1,1,0) with drift	0.27	2.72	2.76	2.07	1.27*	83.01
78	ARIMA(0,1,0) with drift	1.33	1.81	1.12	1.00	0.64*	64.93
79	ARIMA(1,2,0)	0.00	2.83	2.63	2.27	1.00*	66.43
80	ARIMA(1,1,0) with drift	0.47	1.43	2.18	1.71	0.51*	59.76
81	ARIMA(0,1,0) with drift	0.47	1.97	2.03	1.78	0.72*	67.52
82	ARIMA(0,1,0) with drift	0.47	1.26	1.22	1.19	0.85*	76.64
83	ARIMA(1,1,0) with drift	0.41	2.20	2.86	2.31	0.99*	77.94
84	ARIMA(1,1,0) with drift	0.67	1.68	1.91	1.55	0.80*	68.76
85	ARIMA(0,1,1) with drift	0.34	1.49	1.62	1.37	0.81*	64.61
86	ARIMA(1,2,0)	0.03	2.17	2.47	2.04	0.64*	74.76
87	ARIMA(0,1,0) with drift	1.21	3.14	1.89	1.92	1.76*	73.35
88	ARIMA(0,2,1)	0.26	1.83	1.33	1.11	0.92*	64.05
89	ARIMA(0,1,0) with drift	0.89	1.90	1.50	1.47	1.28*	73.86
90	ARIMA(1,1,0)	0.95	0.99	0.82	0.57	0.37*	53.25
91	ARIMA(0,2,1)	0.30	1.46	1.12	0.97	0.51*	60.94
92	ARIMA(0,1,0) with drift	0.40	2.34	2.15	1.80	0.75*	77.70
93	ARIMA(0,1,0) with drift	0.96	3.68	3.02	2.66	1.83*	85.24
94	ARIMA(0,1,1) with drift	0.42	1.18	1.49	1.17	0.62*	59.81
95	ARIMA(0,1,0) with drift	0.56	2.48	1.91	1.71	1.51*	66.27
96	ARIMA(0,2,1)	0.00	3.81	2.39	2.05	1.72*	64.50
97	ARIMA(0,1,0) with drift	0.06	2.15	1.63	1.42	0.68*	64.57
98	ARIMA(1,2,0)	0.17	2.46	1.91	1.61	0.85*	72.25
99	ARIMA(0,1,0) with drift	0.07	2.13	1.62	1.41	0.68*	64.62
100	ARIMA(1,2,0)	0.30	2.29	2.35	2.27	1.25*	62.05
101	ARIMA(0,1,0) with drift	0.89	1.90	1.50	1.47	1.28*	73.86
102	ARIMA(0,1,1) with drift	0.48	1.99	2.22	1.88	0.66*	78.09
103	ARIMA(0,1,1) with drift	0.06	1.44	1.72	1.55	0.78*	61.90
104	ARIMA(0,1,0) with drift	0.85	1.69	1.87	1.84	0.90*	54.73
105	ARIMA(0,2,1)	0.16	4.03	2.83	2.54	2.44*	78.37
106	ARIMA(0,2,0)	0.18	1.95	2.46	1.99	1.03*	68.72
107	ARIMA(1,1,0)	0.95	0.99	0.82	0.57	0.37*	53.25
108	ARIMA(0,1,0) with drift	0.75	0.53	0.29	0.29	0.29*	80.92
109	ARIMA(0,1,0) with drift	0.30	1.77	1.60	1.51	0.67*	56.37
110	ARIMA(0,1,0) with drift	0.03	2.73	2.08	1.64	0.58*	61.38
111	ARIMA(0,2,1)	0.50	3.20	2.73	2.35	1.95*	79.31
112	ARIMA(0,1,0) with drift	0.35	3.64	2.80	2.36	1.88*	70.42
113	ARIMA(0,2,0)	0.19	1.70	2.64	2.18	0.58*	76.34
114	ARIMA(0,1,0) with drift	0.07	2.13	1.62	1.41	0.68*	64.62
115	ARIMA(1,1,0)	0.71	1.85	1.90	1.59*	1.87	55.07
116	ARIMA(0,1,0) with drift	0.06	2.15	1.63	1.42	0.68*	64.57

Continued…

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error				Predicted Weight
			ARIMA with λ	Gomp–ertz	Logistic	Double logistic	
117	ARIMA(1,1,0)	0.71	1.70	2.92	2.49	1.21*	62.79
118	ARIMA(1,1,0) with drift	0.87	1.25	1.49	1.38	0.53*	61.30
119	ARIMA(0,1,0) with drift	0.06	2.15	1.63	1.42	0.68*	64.57
120	ARIMA(0,2,1)	0.34	3.77	2.97	2.85	1.16*	81.79
121	ARIMA(0,2,0)	0.29	1.09	1.82	1.47	0.39*	62.26
122	ARIMA(0,2,0)	-0.08	2.67	2.26	1.81	0.93*	75.62
123	ARIMA(0,1,0) with drift	0.04	2.28	2.09	1.76	0.71*	56.83
124	ARIMA(0,1,0) with drift	0.46	1.69	1.19	0.97	0.80*	62.44
125	ARIMA(0,1,0) with drift	0.25	2.34	1.99	1.62	0.88*	69.30
126	ARIMA(0,1,0) with drift	0.14	3.55	2.17	1.75	1.44*	67.72
127	ARIMA(1,1,0) with drift	0.67	1.68	1.91	1.55	0.80*	68.76
128	ARIMA(1,1,0) with drift	0.41	2.20	2.86	2.31	0.99*	77.94
129	ARIMA(0,1,0) with drift	0.47	1.26	1.22	1.19	0.85*	76.64
130	ARIMA(0,1,0) with drift	0.47	1.97	2.03	1.78	0.72*	67.52
131	ARIMA(1,1,0) with drift	0.47	1.43	2.18	1.71	0.51*	59.76
132	ARIMA(1,2,0)	0.01	2.84	2.62	2.27	1.00*	66.46
133	ARIMA(1,1,0) with drift	0.27	2.72	2.76	2.07	1.27*	83.01
134	ARIMA(1,2,0)	0.27	3.41	3.97	3.51	3.05*	60.62
135	ARIMA(0,1,0) with drift	0.66	1.95	1.95	1.52	0.77*	62.21
136	ARIMA(0,1,0) with drift	0.69	2.51	1.60	1.29	1.18*	72.39
137	ARIMA(1,2,0)	0.21	2.08	1.88	1.28	0.84*	62.90
138	ARIMA(0,2,1)	-0.38	10.35	5.47	5.44	1.39*	108.01
139	ARIMA(0,1,1) with drift	0.34	2.51	2.15	1.71	1.30*	70.19
140	ARIMA(1,1,0) with drift	0.46	3.39	2.72	2.71	1.88*	62.95
141	ARIMA(0,2,1)	0.26	1.83	1.33	1.11	0.92*	64.05
142	ARIMA(1,1,0) with drift	0.65	1.42	2.09	1.89	0.63*	66.49
143	ARIMA(0,1,0) with drift	0.30	1.93	2.05	1.89	0.70*	55.92
144	ARIMA(1,1,0) with drift	0.19	2.29	2.66	2.19	0.91*	69.85
145	ARIMA(0,1,0) with drift	0.77	1.23	0.99	0.93	0.78*	65.81
146	ARIMA(0,2,0)	0.38	1.83	2.69	2.14	1.37*	63.89
147	ARIMA(0,2,1)	0.29	6.21	5.27	4.77	4.28*	81.28
148	ARIMA(0,2,0)	-0.08	2.97	2.84	2.38	0.72*	59.77
149	ARIMA(0,2,0)	0.53	1.96	1.68	1.37	1.23*	65.37
150	ARIMA(0,1,0) with drift	0.23	2.14	1.79	1.53	0.70*	65.50
151	ARIMA(1,2,0)	-0.19	4.00	2.28	2.18	2.05*	73.66
152	ARIMA(1,1,0) with drift	0.32	2.00	2.55	2.33	0.62*	61.93
153	ARIMA(0,1,0) with drift	0.58	3.02	1.82	1.80	1.69*	82.30
154	ARIMA(0,1,0) with drift	0.12	1.79	1.33	1.14	0.60*	106.73
155	ARIMA(1,2,0)	0.37	1.17	1.06	0.71	0.46*	56.74
156	ARIMA(0,1,1)	1.03	4.25	3.23	3.16	2.64*	56.85
157	ARIMA(1,1,0)	0.65	0.99	1.76	1.40	0.42*	59.37
158	ARIMA(1,1,0) with drift	0.61	1.75	2.42	2.08	0.61*	69.42
159	ARIMA(2,1,0)	0.40	0.80*	2.18	1.97	0.90	50.81
160	ARIMA(1,2,0)	0.21	1.57	1.82	1.55	0.68*	53.65
161	ARIMA(0,1,0) with drift	0.69	2.42	2.10	1.95	1.05*	84.55
162	ARIMA(0,1,0) with drift	0.34	2.50	2.65	2.43	1.02*	64.22
163	ARIMA(0,1,0) with drift	0.37	3.65	2.43	2.13	1.93*	84.58
164	ARIMA(0,2,0)	0.28	1.98	2.52	2.04	0.72*	68.70
165	ARIMA(0,1,0) with drift	0.49	2.12	1.73	1.59	1.00*	63.24

Continued…

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error				Predicted Weight
			ARIMA with λ	Gomp–ertz	Logistic	Double logistic	
166	ARIMA(1,2,0)	0.41	1.31	1.33	1.03	0.42*	63.23
167	ARIMA(0,1,0) with drift	-0.02	2.80	1.94	1.67	0.93*	92.78
168	ARIMA(0,1,0) with drift	0.44	1.84	1.48	1.32	0.95*	78.89
169	ARIMA(0,1,0) with drift	1.10	1.75	1.31	1.30	0.63*	90.98
170	ARIMA(1,1,0) with drift	0.07	1.73	2.27	1.79	0.44*	73.00
171	ARIMA(0,1,0) with drift	0.33	2.95	2.38	2.24	1.61*	80.20
172	ARIMA(0,1,0) with drift	0.00	1.78	1.25	1.25	0.45*	73.51
173	ARIMA(0,1,0) with drift	0.92	2.93	2.25	1.90	1.39*	77.05
174	ARIMA(0,1,0) with drift	0.62	1.84	1.90	1.56	0.83*	58.81
175	ARIMA(0,1,0) with drift	0.72	2.20	2.13	1.77	0.97*	59.13
176	ARIMA(0,1,0) with drift	0.09	2.08	2.25	1.33	0.52*	72.41
177	ARIMA(0,2,1)	-0.04	2.43	1.71	1.35	0.83*	68.45
178	ARIMA(0,1,0) with drift	0.34	2.18	1.94	1.58	0.81*	111.48
179	ARIMA(0,1,0) with drift	0.31	1.59	1.54	1.53	0.88*	68.66
180	ARIMA(0,1,0) with drift	0.73	1.65	1.64	1.54	0.83*	79.95

Note: The asterisk * indicates minimum MSE and selected the model and D.S. No. means Data Set Number

Table 5.4b Weight estimation using ARIMA model with Box-Cox transformation, Gompertz, logistic and double logistic models for girls

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error				Predicted Weight
			ARIMA with λ	Gomp–ertz	Logistic	Double logistic	
01	ARIMA(0,1,0) with drift	1.41	1.61	1.61	1.79	0.72*	47.26
02	ARIMA(0,2,0)	0.08	1.90	1.48	1.11	0.51*	50.57
03	ARIMA(0,2,0)	0.18	2.15	1.83	1.61	0.62*	53.79
04	ARIMA(0,1,0) with drift	1.29	1.58	1.55	1.39	0.72*	48.73
05	ARIMA(0,1,0) with drift	0.69	2.14	1.85	1.58	1.00*	57.00
06	ARIMA(0,1,0) with drift	0.84	2.17	1.42	18.51	1.12*	60.59
07	ARIMA(0,1,1) with drift	0.99	1.18	1.65	1.49	0.68*	56.23
08	ARIMA(1,2,0)	0.37	1.81	2.25	1.89	1.12*	50.85
09	ARIMA(0,2,0)	0.44	1.81	2.59	2.17	0.75*	53.74
10	ARIMA(0,1,0) with drift	0.83	2.05	1.81	1.62	0.87*	50.92
11	ARIMA(0,2,0)	0.00	2.58	2.42	1.81	0.74*	63.07
12	ARIMA(2,2,0)	-0.02	2.31	1.70	1.39	1.01*	48.41
13	ARIMA(1,1,0) with drift	2.06	1.81	2.14	1.91	0.83*	46.27
14	ARIMA(0,1,1) with drift	1.86	0.99	0.63	0.62	0.57*	45.23
15	ARIMA(1,2,0)	-0.02	1.97	1.95	1.52	0.84*	49.50
16	ARIMA(0,2,1)	0.42	1.96	1.69	1.45	0.83*	53.60
17	ARIMA(1,2,0)	-0.02	1.97	1.95	1.52	0.84*	49.50
18	ARIMA(0,1,1) with drift	1.86	0.99	0.63	0.62	0.57*	45.23
19	ARIMA(0,2,1)	0.43	1.96	1.70	1.45	1.06*	51.02
20	ARIMA(1,2,0)	0.62	2.01	3.17	2.70	1.77*	51.23
21	ARIMA(0,2,0)	0.78	1.46	1.77	1.46	0.53*	48.18
22	ARIMA(1,1,0)	0.08	2.23	2.42	2.07	1.49*	51.78
23	ARIMA(1,2,0)	0.23	1.58	1.57	1.26	0.43*	56.38
24	ARIMA(0,1,0) with drift	0.33	2.44	2.22	1.84	0.73*	60.28

Continued…

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error ARIMA with λ	Gomp–ertz	Logistic	Double logistic	Predicted Weight
25	ARIMA(0,2,1)	0.74	1.51	1.84	1.60	0.62*	45.30
26	ARIMA(1,2,0)	0.56	1.71	1.71	1.47	0.95*	45.96
27	ARIMA(1,1,0) with drift	0.38	1.38	1.49	1.28	0.75*	54.60
28	ARIMA(0,1,1) with drift	-0.09	1.85	2.00	1.66	0.63*	49.72
29	ARIMA(0,1,1) with drift	0.96	1.41	1.60	1.43	0.51*	50.60
30	ARIMA(0,2,0)	0.65	2.19	2.08	1.82	1.13*	50.30
31	ARIMA(1,1,0) with drift	0.29	2.84	2.91	2.43	1.48*	55.48
32	ARIMA(0,1,0) with drift	0.64	1.19	0.89	0.80	0.52*	47.93
33	ARIMA(0,1,0) with drift	0.66	2.74	2.39	2.00	1.57*	55.86
34	ARIMA(0,1,0) with drift	0.97	1.26	1.39	1.30	0.56*	46.60
35	ARIMA(0,1,0) with drift	0.31	2.12	1.35	1.13	0.82*	45.45
36	ARIMA(0,2,0)	1.08	1.13	1.03	0.74	0.50*	46.93
37	ARIMA(1,2,0)	0.62	1.52	1.22	0.89	0.76*	55.94
38	ARIMA(0,1,0) with drift	0.60	1.83	1.44	1.22	0.97*	50.66
39	ARIMA(0,2,0)	0.18	1.93	1.82	1.30	0.59*	54.83
40	ARIMA(0,1,0) with drift	0.66	1.38	1.23	1.05	0.67*	60.39
41	ARIMA(0,2,0)	0.52	2.66	2.73	2.31	1.12*	52.77
42	ARIMA(0,2,1)	0.09	1.52	1.17	0.96	0.55*	41.55
43	ARIMA(0,2,0)	0.46	1.44	1.97	1.73	0.51*	57.23
44	ARIMA(0,1,0) with drift	2.45	2.24	2.40	2.40	1.09*	46.98
45	ARIMA(0,1,0) with drift	1.28	1.30	1.16	1.02	0.80*	51.45
46	ARIMA(0,2,2)	0.37	3.87	2.35*	2.44	2.39	60.43
47	ARIMA(0,2,0)	-0.35	1.66	1.12	0.70*	0.90	48.73
48	ARIMA(0,2,0)	0.63	1.07	1.62	1.37	0.38*	43.52
49	ARIMA(0,2,1)	0.44	1.68	1.41	1.33	0.43*	49.97
50	ARIMA(2,2,0)	0.11	1.54	2.28	1.89	0.88*	53.62
51	ARIMA(0,2,0)	0.14	2.27	1.65	1.23	0.98*	54.98
52	ARIMA(1,1,0) with drift	0.91	1.21	1.10	0.75	0.49*	46.31
53	ARIMA(0,1,0) with drift	1.08	2.91	2.12	1.76	1.49*	68.27
54	ARIMA(1,2,0)	-0.09	1.24	0.93	0.77	0.45*	48.80
55	ARIMA(1,1,0)	0.70	1.47	1.73	1.51	0.59*	52.47
56	ARIMA(0,2,0)	0.01	1.86	2.24	1.98	0.84*	55.33
57	ARIMA(0,1,0) with drift	0.95	1.92	1.23	0.99	0.90*	61.19
58	ARIMA(1,2,0)	0.36	1.93	1.85	1.42	1.17*	52.69
59	ARIMA(1,1,0) with drift	0.55	1.69	2.35	1.94	0.52*	54.95
60	ARIMA(0,2,1)	0.99	1.98	1.79	1.54	0.91*	44.54
61	ARIMA(0,1,1) with drift	0.37	1.55	1.74	1.46	0.74*	52.61
62	ARIMA(0,2,1)	0.16	1.72	1.28	0.97	0.58*	52.13
63	ARIMA(0,1,0) with drift	1.15	1.98	1.55	1.30	0.63*	49.43
64	ARIMA(1,1,0) with drift	0.74	0.98	1.29	1.09	0.54*	49.09
65	ARIMA(0,2,0)	0.26	1.60	1.46	1.11	0.46*	57.96
66	ARIMA(0,1,0) with drift	0.49	3.03	2.28	1.86	1.57*	55.96
67	ARIMA(1,1,0)	1.38	1.29	1.52	1.32	0.42*	44.91
68	ARIMA(0,2,1)	0.63	1.86	1.72	1.49	1.01*	46.74
69	ARIMA(1,2,0)	0.07	1.71	1.99	1.68	0.60*	48.19
70	ARIMA(1,2,0)	0.41	1.75	1.91	1.52	0.62*	52.26
71	ARIMA(1,2,0)	0.78	1.37	1.10	0.85	0.48*	48.69
72	ARIMA(0,1,0) with drift	0.85	0.81	0.60	0.58	0.37*	53.18

Continued…

D.S. No.	ARIMA Model	Box-Cox λ	Root Mean Square Error				Predicted Weight
			ARIMA with λ	Gomp–ertz	Logistic	Double logistic	
73	ARIMA(1,2,0)	0.98	1.33	1.12	0.89	0.52*	58.57
74	ARIMA(0,1,0) with drift	0.96	0.99	0.87	0.82	0.55*	44.28
75	ARIMA(0,1,0) with drift	1.35	1.91	1.32	1.07	0.84*	53.63
76	ARIMA(0,1,0) with drift	1.86	2.14	1.45	1.22	1.06*	48.48
77	ARIMA(1,2,0)	0.52	1.50	1.77	1.37	0.53*	51.22
78	ARIMA(0,2,0)	0.16	1.85	1.86	1.46	0.72*	41.39
79	ARIMA(0,1,0) with drift	0.95	2.51	1.65	1.43	1.27*	61.35
80	ARIMA(0,2,0)	0.60	2.03	2.00	1.64	0.96*	52.23
81	ARIMA(0,2,1)	0.14	1.31	2.09	1.54	0.42*	51.17
82	ARIMA(0,2,1)	-0.32	2.57	1.80	1.37*	2.26	54.62
83	ARIMA(0,1,0) with drift	0.14	3.29	2.61	2.53	1.65*	60.96
84	ARIMA(0,1,0) with drift	0.80	1.92	1.77	1.56	0.93*	60.07
85	ARIMA(0,2,0)	0.62	2.00	2.71	2.38	0.62*	53.38
86	ARIMA(0,2,0)	0.31	1.38	2.16	1.86	0.63*	53.26
87	ARIMA(0,1,0) with drift	1.06	1.06	0.79	0.76	0.36*	55.31
88	ARIMA(0,2,0)	0.93	1.46	1.84	1.66	0.50*	50.68
89	ARIMA(0,1,0) with drift	1.11	2.21	1.62	1.47	1.21*	55.46
90	ARIMA(0,1,0) with drift	0.53	2.60	1.83	1.78	1.41*	73.92
91	ARIMA(1,2,0)	0.15	2.77	2.10	1.77	1.27*	47.81
92	ARIMA(1,2,0)	0.65	1.42	1.86	1.41	1.05*	45.19
93	ARIMA(0,1,0) with drift	-0.12	3.47	2.09	2.08	1.36*	50.14
94	ARIMA(1,1,0) with drift	0.64	1.40	1.63	1.36	0.61*	47.06
95	ARIMA(0,2,0)	0.45	1.44	2.23	1.79	1.00*	49.72
96	ARIMA(0,2,1)	-0.18	2.78	1.39	1.31	1.17*	42.07
97	ARIMA(0,1,0) with drift	0.83	1.88	1.10	1.06	0.87*	58.64
98	ARIMA(0,1,0) with drift	1.38	1.64	1.68	1.49	0.71*	49.29
99	ARIMA(0,1,0) with drift	1.45	2.19	1.82	1.51	1.19*	56.30
100	ARIMA(0,1,0) with drift	0.42	4.09	2.48	2.36	2.07*	67.46
101	ARIMA(0,2,0)	0.40	1.51	2.08	1.64	0.77*	52.61
102	ARIMA(0,2,1)	0.56	2.78	3.15	2.84	1.66*	47.94
103	ARIMA(0,1,1) with drift	0.94	0.58	0.67	0.76	0.45*	45.68
104	ARIMA(1,2,0)	0.60	1.44	1.11	0.80	0.51*	45.93
105	ARIMA(1,2,0)	0.00	3.12	2.49	1.95	1.03*	50.95
106	ARIMA(0,1,0) with drift	0.44	1.77	1.31	1.17	1.03*	47.48
107	ARIMA(0,1,0) with drift	0.52	2.16	1.80	1.69	1.19*	54.36
108	ARIMA(0,1,0) with drift	0.66	1.50	1.40	1.18	0.49*	58.37
109	ARIMA(0,2,0)	0.69	1.33	1.81	1.56	0.53*	46.23
110	ARIMA(0,1,0) with drift	0.78	1.73	1.68	1.46	0.81*	56.59
111	ARIMA(0,2,1)	0.78	2.59	2.67	2.31	1.18*	53.28
112	ARIMA(1,1,0) with drift	0.61	1.06	1.25	0.89	0.44*	48.41
113	ARIMA(0,2,0)	1.29	1.06	1.29	1.28	0.42*	57.04
114	ARIMA(0,1,0) with drift	0.41	2.41	2.31	1.92	1.10*	48.97
115	ARIMA(0,2,0)	0.23	2.33	1.98	1.69	0.67*	53.91
116	ARIMA(1,1,0)	0.06	2.23	2.42	2.07	1.48*	51.77
117	ARIMA(2,2,0)	0.67	1.73*	3.16	2.69	1.75	46.33
118	ARIMA(0,1,0) with drift	1.05	1.09	0.95	0.77	0.39*	46.82
119	ARIMA(0,2,1)	0.74	1.51	1.66	1.60	0.62*	45.30
120	ARIMA(0,1,0) with drift	0.33	2.44	2.22	1.84	0.73*	60.28

Note: The asterisk * indicates minimum MSE and selected the model and D.S. No. means data set number

5.6.2 Second Stage

In this stage, final stature was estimated from the proposed model. The value of age at the final stature was 25 and that of weight at age 25 was obtained from Tables 5.4a and 5.4b for boys and girls, respectively. Using the estimated proposed model for each data set with corresponding age equal 25 and estimated weight at age 25, the value of estimated final stature could be obtained. The predicted (estimated) final statures for boys and girls were presented in Table 5.5a and Table 5.5b, respectively.

Table 5.5a Predicted final stature for boys

D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature
01	170.62	**37**	168.24	**73**	168.52	**109**	170.55	**145**	172.21
02	171.58	**38**	175.67	**74**	172.56	**110**	176.98	**146**	168.25
03	173.42	**39**	179.32	**75**	175.29	**111**	171.20	**147**	172.08
04	182.99	**40**	165.53	**76**	170.31	**112**	185.10	**148**	170.22
05	175.18	**41**	173.37	**77**	180.39	**113**	174.76	**149**	170.68
06	170.79	**42**	170.11	**78**	174.53	**114**	168.26	**150**	172.34
07	170.93	**43**	173.78	**79**	168.70	**115**	161.16	**151**	165.39
08	173.89	**44**	170.07	**80**	169.05	**116**	168.18	**152**	173.44
09	166.30	**45**	172.74	**81**	176.76	**117**	176.00	**153**	169.50
10	170.13	**46**	172.42	**82**	173.39	**118**	163.65	**154**	175.19
11	176.78	**47**	164.64	**83**	178.53	**119**	168.18	**155**	169.20
12	175.22	**48**	171.77	**84**	162.93	**120**	177.02	**156**	172.99
13	171.62	**49**	167.10	**85**	177.17	**121**	181.34	**157**	173.58
14	173.95	**50**	173.14	**86**	183.23	**122**	179.94	**158**	178.14
15	180.31	**51**	180.39	**87**	164.71	**123**	173.04	**159**	163.49
16	175.72	**52**	172.33	**88**	168.33	**124**	170.12	**160**	174.48
17	166.70	**53**	180.22	**89**	176.12	**125**	170.86	**161**	170.08
18	170.33	**54**	170.04	**90**	175.13	**126**	173.70	**162**	172.07
19	168.60	**55**	171.99	**91**	172.01	**127**	162.93	**163**	178.38
20	168.90	**56**	180.38	**92**	177.01	**128**	178.63	**164**	182.35
21	177.29	**57**	182.35	**93**	185.02	**129**	173.39	**165**	168.88
22	166.65	**58**	168.88	**94**	171.71	**130**	176.16	**166**	180.89
23	169.55	**59**	165.91	**95**	172.60	**131**	169.03	**167**	178.42
24	169.45	**60**	168.58	**96**	178.14	**132**	169.44	**168**	173.39
25	170.04	**61**	175.51	**97**	168.18	**133**	180.41	**169**	170.49
26	170.56	**62**	175.60	**98**	176.60	**134**	170.18	**170**	180.37
27	169.04	**63**	176.90	**99**	168.26	**135**	169.22	**171**	182.62

Continued…

D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature
28	171.24	**64**	164.14	**100**	166.94	**136**	172.56	**172**	172.51
29	173.92	**65**	184.63	**101**	176.12	**137**	175.29	**173**	171.31
30	175.94	**66**	177.35	**102**	177.33	**138**	172.74	**174**	175.63
31	177.89	**67**	175.63	**103**	161.29	**139**	175.99	**175**	168.61
32	167.46	**68**	168.28	**104**	168.04	**140**	170.44	**176**	169.10
33	176.64	**69**	174.48	**105**	184.95	**141**	168.33	**177**	164.14
34	172.21	**70**	178.03	**106**	176.63	**142**	168.45	**178**	183.27
35	168.45	**71**	166.00	**107**	175.13	**143**	165.16	**179**	177.22
36	173.45	**72**	168.26	**108**	174.85	**144**	174.40	**180**	175.39

Note: D.S. No. means Data Set Number

Table 5.5b Predicted final stature for girls

D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature	D.S. No.	Final Stature
01	157.67	**25**	166.02	**49**	166.23	**73**	165.32	**97**	158.82
02	164.22	**26**	150.59	**50**	162.73	**74**	154.38	**98**	162.29
03	157.63	**27**	165.86	**51**	156.14	**75**	165.82	**99**	162.50
04	160.28	**28**	158.05	**52**	153.94	**76**	156.88	**100**	160.83
05	163.80	**29**	162.63	**53**	161.50	**77**	159.88	**101**	158.29
06	158.76	**30**	161.16	**54**	154.54	**78**	157.92	**102**	158.03
07	168.37	**31**	157.25	**55**	160.50	**79**	160.30	**103**	158.77
08	164.15	**32**	162.14	**56**	166.26	**80**	163.52	**104**	151.07
09	150.34	**33**	150.42	**57**	156.20	**81**	161.05	**105**	154.25
10	159.82	**34**	165.06	**58**	159.03	**82**	162.01	**106**	154.37
11	160.38	**35**	156.78	**59**	154.43	**83**	163.41	**107**	157.94
12	156.77	**36**	155.76	**60**	151.74	**84**	161.60	**108**	159.54
13	154.49	**37**	165.97	**61**	155.41	**85**	162.42	**109**	154.13
14	157.96	**38**	162.22	**62**	159.32	**86**	158.75	**110**	160.08
15	157.43	**39**	162.60	**63**	152.14	**87**	157.34	**111**	155.26
16	162.31	**40**	165.29	**64**	159.05	**88**	158.16	**112**	154.92
17	157.43	**41**	165.36	**65**	160.52	**89**	163.42	**113**	157.12
18	157.96	**42**	148.21	**66**	151.47	**90**	161.64	**114**	157.31
19	161.24	**43**	164.51	**67**	152.84	**91**	151.42	**115**	164.60
20	160.60	**44**	157.99	**68**	153.88	**92**	148.93	**116**	156.98
21	155.37	**45**	153.53	**69**	160.81	**93**	163.64	**117**	159.41
22	157.34	**46**	152.59	**70**	153.93	**94**	155.08	**118**	160.14
23	161.36	**47**	158.79	**71**	162.11	**95**	153.30	**119**	166.02
24	169.56	**48**	157.82	**72**	157.66	**96**	157.75	**120**	164.68

Note: D.S. No. means Data Set Number

5.7 Building Equations for Predicting Final Stature

The proposed model was run on the individual longitudinal data of stature to find out the distance curve for each individual data set. Predicted stature at age 25, from age 2 to 13 years and parents statures were considered here for further analysis to build up some prediction equations. Stature at age 2 (S_2), stature at the age 3 (S_3), stature at the age 4 (S_4), stature at the age 5 (S_5), stature at the age 6 (S_6), stature at the age 7 (S_7), stature at the age 8 (S_8), stature at the age 9 (S_9), stature at the age 10 (S_{10}), stature at the age 11 (S_{11}), stature at the age 12 (S_{12}), stature at the age 13 (S_{13}), father stature (*FS*), mother stature (*MS*) and predicted final stature (*PFS*) were considered to build up some equations using selected important variables. At first, multicollinearity problems among the regressors could be checked.

5.7.1 Checking Multicollinearity

The tolerance and variance inflection function (VIF) presented in Table 5.6 that were used to identify the problem of multicollinearity in the regressors. Table 5.6 showed that the calculated tolerance for both boys and girls were some close to zero and some close to one. The tolerance clearly indicated the multicollinearity problem exists in the data set. The maximum values of VIF for boys and girls were 81.198 and 65.755, respectively, which were exceeded 5. Therefore, tolerance and variance inflection function were clearly indicated that the multicollinearity problem exists.

Table 5.6 Collinearity statistics of tolerance and VIF for boys and girls

Dependent Variable	Boys			Girls		
	R^2 value	Tolerance	VIF	R^2 value	Tolerance	VIF
S_2	0.674	0.326	3.064	0.678	0.322	3.110
S_3	0.802	0.198	5.042	0.792	0.208	4.800
S_4	0.890	0.110	9.123	0.847	0.153	6.540
S_5	0.908	0.092	10.824	0.896	0.104	9.587
S_6	0.922	0.078	12.836	0.902	0.098	10.166
S_7	0.971	0.029	33.989	0.967	0.033	30.024
S_8	0.986	0.014	71.448	0.985	0.015	65.748
S_9	0.987	0.013	76.778	0.985	0.015	64.812
S_{10}	0.984	0.016	61.594	0.980	0.020	49.208
S_{11}	0.984	0.016	61.803	0.969	0.031	32.644
S_{12}	0.976	0.024	41.358	0.972	0.028	36.073
S_{13}	0.922	0.078	12.775	0.931	0.069	14.497
FS	0.082	0.918	1.089	0.219	0.781	1.280
MS	0.125	0.875	1.142	0.224	0.776	1.289

5.7.2 Forward Stepwise Ridge Regression Model

The tolerance and variance inflection function showed that the multicollinearity problem was present in the regressors (Table 5.6). The aim of the present phases was to select some important variables that variables could able to predict the final stature. These important variables could be selected by using forward stepwise ridge regression model. Although, stepwise regression model do not much affect by multicollinearity. The challenging issue was to select appropriate value of ridge regression parameter (say, k). There were many methods available to choose the optimum value of k. The modified HKB estimator (Hoerl et al., 1975) was used to select the value of k using the statistical software R i386 3.0.1 with the package 'MASS'. The appropriate values of k were found as 0.0037 and 0.0010 for boys and girls, respectively. Forward stepwise ridge regression model estimated the parameters and the important variables, by using the optimum value of k and STATISTICA 8.0 software. The full summary statistics of forward stepwise ridge regression model for boys and girls were shown in Appendix-3 (Tables 1 and 3).

And, a summary of selected step of the forward stepwise ridge regression model and previews investigation (Ali and Ohtsuki, 2001; Rahman and Ali, 2003) for the dependent variable *PFS* based on stature-variables were shown in Table 5.7a and 5.7b for boys and girls, respectively. The regression coefficients were highly significant with high values of R^2 as well as adjusted R^2 and the standard errors of the estimate attained smaller amount (Table 5.7a). This Table also exhibited that three–variables regression equation which indicated that the 'parent statures' and 'stature at the age 9' were needed to predict the final stature of boys. Similarly, Table 5.7b showed the regression's coefficients were highly significant and standard errors of this estimate attained smaller amount. 'Parent statures' and 'stature at age 13' were needed to predict the final stature of girls.

Finally, the proposed equations of predicting the final stature for the Japanese boys and girls were as follows:

For boys: $PFS = 0.333572\ MS + 0.342368\ S_9 + 0.322908\ FS$

For girls: $PFS = 0.387948\ S_{13} + 0.316876\ MS + 0.294889\ FS$

Table 5.7a Summary of forward stepwise ridge regression model for the dependent variable *PFS* based on stature-variables for boys

Sample Size	Step + in	Variable	Coefficient (*p*-value)	Standard error	F–to inter /remove (*p*-value)	R^2 (Adj. R^2) [DW Value]
			Present investigation			
180	1	*MS*	0.333572 (0.000001)	0.038531	34039.71	0.99926973
	2	S_9	0.342368 (0.000001)	0.036509	(0.000001)	(0.99824040)
	3	*FS*	0.322908 (0.000001)	0.037915		[2.0213]
			Ali and Ohtsuki(2001) investigation			
410	1	S_9	1.270660	0.096113	377603.7	0.99892
	2	S_3	0.638750	0.079237		
	3	S_{12}	-0.377094	0.072787		
			Rahman and Ali (2003) investigation			
464	1	S_9	1.046828	0.045456	548059.2	0.999156
	2	S_3	0.397943	0.063855		

Table 5.7b Summary of forward stepwise ridge regression model for the dependent variable *PFS* based on stature-variables for girls

Sample Size	Step + in	Variable	Coefficient (*p*-value)	Standard error	F–to inter /remove (*p*-value)	R^2 (Adj. R^2) [DW Value]
			Present investigation			
120	1	S_{13}	0.387948 (0.000001)	0.051731	50900.09	0.99973438
	2	*MS*	0.316876 (0.000001)	0.047571	(0.000001)	(0.99921475)
	3	*FS*	0.294889 (0.000001)	0.050045		[1.6557]
			Ali and Ohtsuki(2001) investigation			
262	1	S_{13}	3.20074	0.143402	393773.6	0.99934
	2	S_{12}	-3.29566	0.280664		
	3	S_{11}	1.11848	0.155298		
			Rahman and Ali (2003) investigation			
259	1	S_{13}	2.87668	0.102438	345503.1	0.999254
	2	S_{12}	-1.89627	0.105918		

Diagnostic Checking: Although, the forward stepwise ridge regression model considered the variable(s) at each step when the variable(s) was significant. But the other tested such as line graph of predicted *vs.* observed values, outlier, influential data point, histogram of residuals, autocorrelation of residuals, and normal probability plot were must be checked for the validity of forward stepwise ridge regression model. The standardized residuals greater than 3 in absolute value were considered as outlier. Also, the Cook's distance was greater than one, indicating the influential observation. The observed, predicted, residual, standard predicted, standardized residual, standard error of predicted and Cook's distance values of the predicted final stature of boys and girls were shown in Appendix-3 (Tables 2 and 4). Table 2 (Appendix-3) showed there were no points of absolute values of standardized residuals and Cook's distance falls outside of 3 and greater than 1, respectively. Therefore, the predicting equation of boys was free from outlier and influential data point. The line graph of predict *vs.* observed value (Figure 5.4a), histogram of residuals (Figure 5.4b) and normal probability plot (Figure 5.4c) of residuals showed that the fit was well and residuals were approximately normally distributed. The estimated Durbin–Watson *d* value equal to 2.0213 (Table 5.7a), indicated no positive or negative autocorrelation in the residuals. Similarly, from Table 4 (Appendix-3), Figure 5.5a, Figure 5.5b, and Figure 5.5c, there were no outlier and influential points and the fit was also well and residuals were normally distributed. The estimated Durbin–Watson *d* value was 1.6557 (Table 5.7b), which was fall in zone of in–conclusion region.

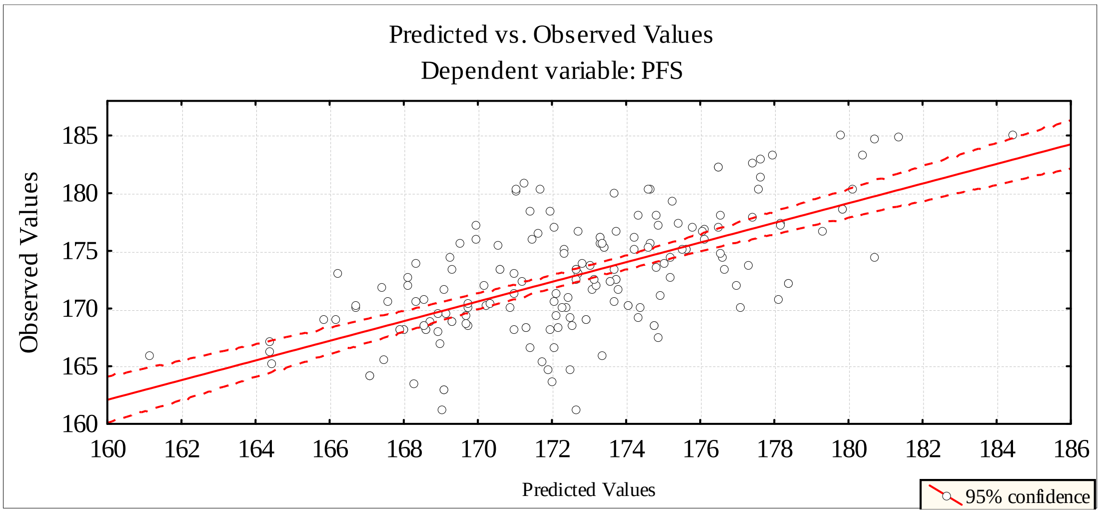


Figure 5.4a Predicted *vs.* observed values for predicting equation of boys

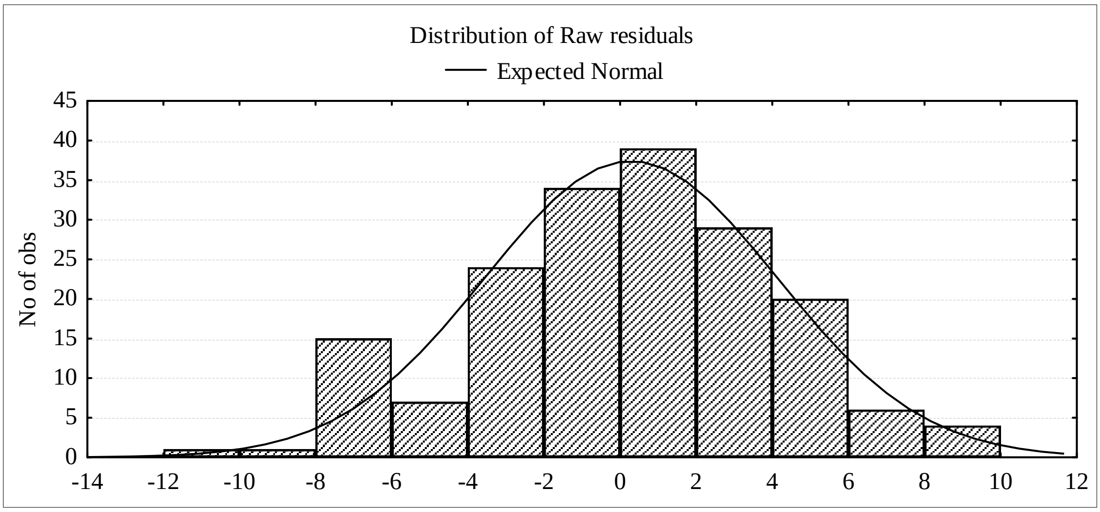


Figure 5.4b Histogram of residuals for predicting equation of boys

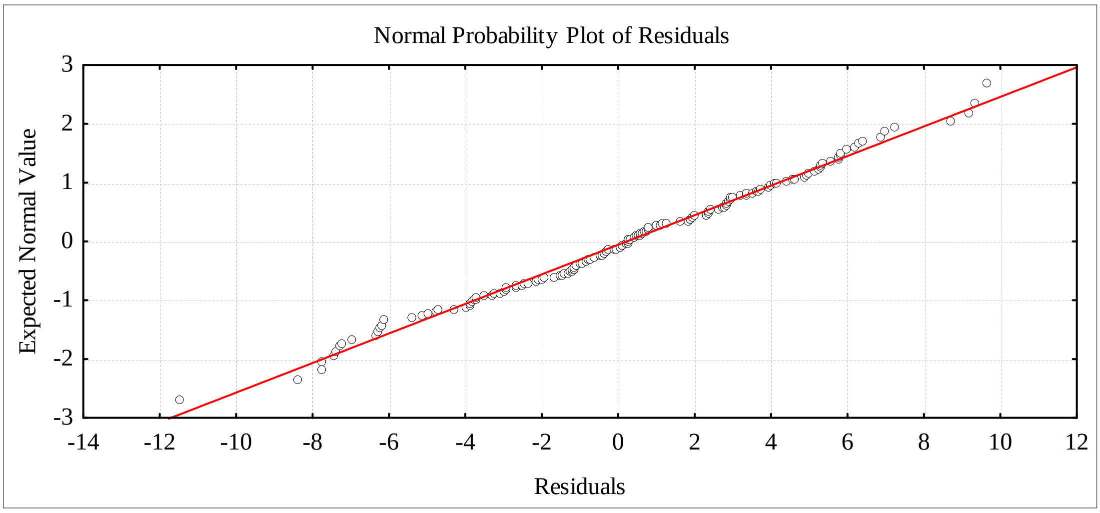


Figure 5.4c Normal probability plot of residuals for predicting equation of boys

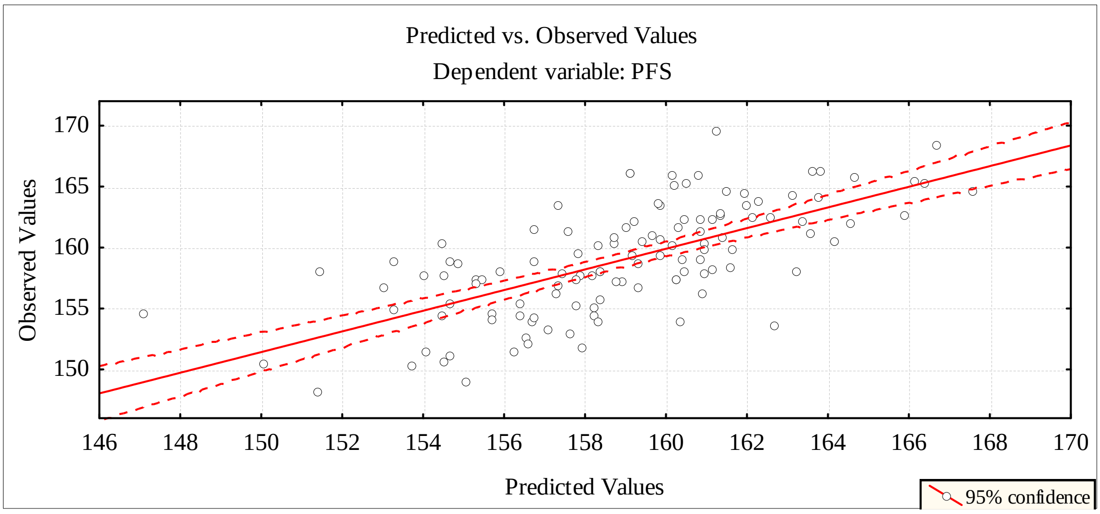


Figure 5.5a Predicted *vs.* observed values for predicting equation of girls

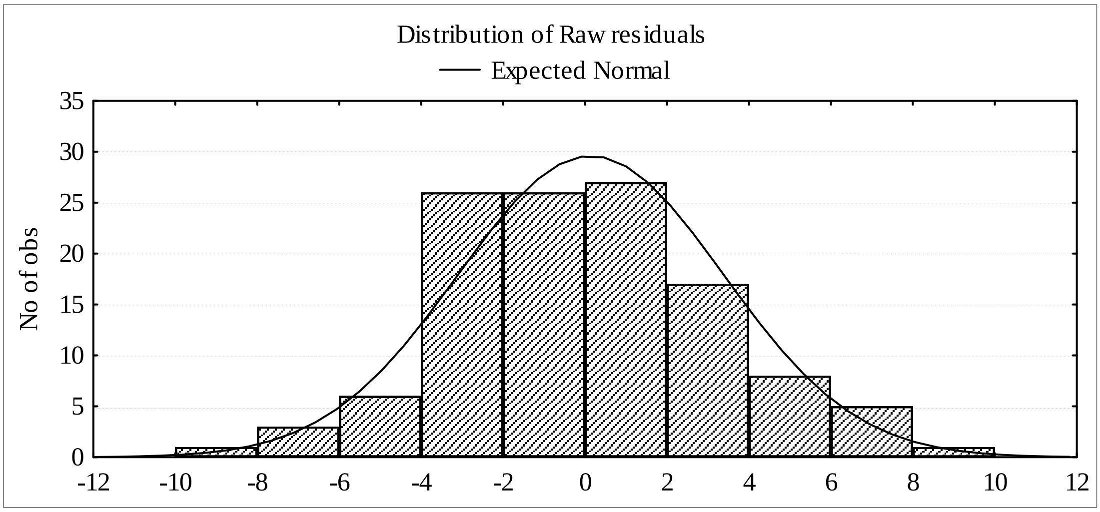


Figure 5.5b Histogram of residuals for predicting equation of girls

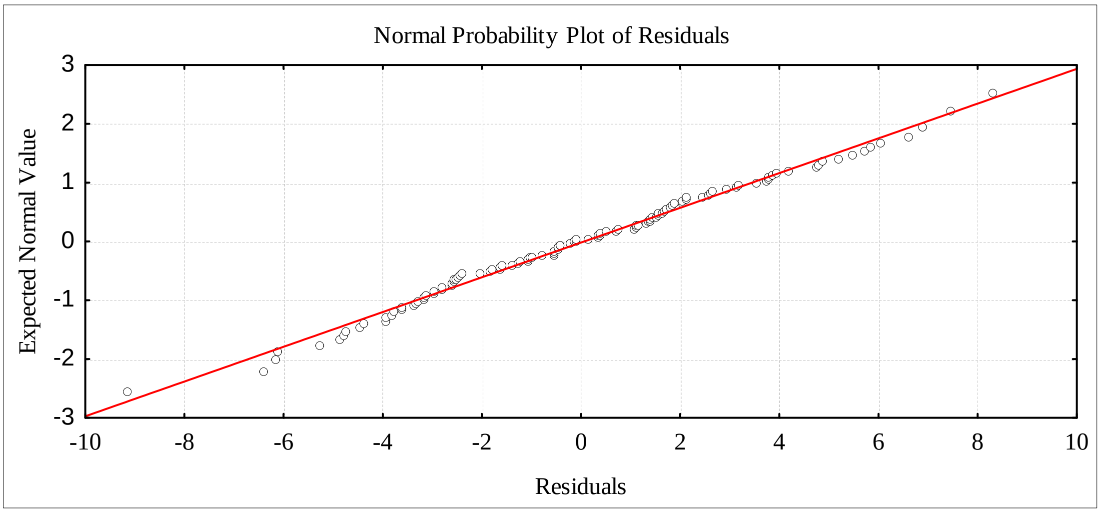


Figure 5.5c Normal probability plot of residuals for predicting equation of girls

Precision of the Estimated Model: Analysis of residual for boys and girls were considered to understand the precision of the prediction for final stature. Average of observed, predicted, residual, 95% confidence bounds of mean residuals and standard error (SE) of predicted equations of final stature for boys and girls were shown in Table 5.8. The average value of residual was 0.091 cm and standard error of residuals and prediction were as 3.93 cm and 0.28 cm, respectively for boys that were small, implying the prediction equation was sufficient for predicting the final stature. Similarly, the average value of residual was 0.023 cm and standard error of residuals and prediction were as 3.33 cm and 0.23 cm, respectively for girls that were also small, implying the prediction equation was sufficient for predicting the final stature. Therefore, the prediction of final stature based on different stature variables was, on average, under-estimated by 0.091 cm and 0.023 cm for boys and girls, respectively.

Table 5.8 Average of observed, predicted, residual, 90% confidence bounds of residuals and standard error (SE) of predicted equations of final stature for boys and girls. Values in the first bracket ( ) were of Ali and Ohtsuki (2001) and that in third bracket [ ] were of Rahman and Ali (2003)

Data Set size	**Observed stature (cm)**	**Predicted stature (cm)**	**Residual (cm)**		**95% confidence bound of mean residuals (cm)**		**SE of prediction (cm)**
			Mean	**SD**	**Lower**	**Upper**	
			Boys				
180	172.85	172.76	0.091	3.93	-0.35	0.80	0.28
(410)	(172.34)	(172.23)	(0.113)	(5.17)	(-6.024)	(7.497)	(0.42)
[464]	[171.69]	[171.58]	[0.109]	[4.80]	[-0.258]	[0.476]	[0.30]
			Girls				
120	158.96	158.94	0.023	3.33	-0.54	0.67	0.23
(262)	(159.00)	(159.19)	(0.027)	(2.80)	(-2.652)	(3.512)	(0.29)
[259]	[158.44]	[158.39]	[0.046]	[2.89]	[-0.250]	[0.344]	[0.24]

Cross Validity Predictive Power: The proposed predicted equations to predict the final stature of Japanese boys and girls were cross validated by the cross validity predictive power. Estimated cross validity predictive power of the predicted equations of Japanese boys and girls were shown in Table 5.9. Table 5.9 showed, for any independent sample of the Japanese population more than 99% of the variance on the predicted variable, *PFS*, would be explained by the proposed equations for both boys and girls. Therefore, the predicted equations for boys and girls were highly cross validated.

Table 5.9 Estimated cross validity predictive power of the predicted equations for boys and girls. Values in the first bracket ( ) were of Ali and Ohtsuki (2001) and that in third bracket [ ] were of Rahman and Ali (2003)

$\boldsymbol{N}$	$\boldsymbol{P}$	$\boldsymbol{R^2}$	$\boldsymbol{\rho_{cv}^2}$
		Boys	
180	3	0.99926973	0.999240353
(410)	(3)	(0.99910)	(0.99908)
[464]	[2]	[0.99922]	[0.99921]
		Girls	
120	3	0.99973438	0.999718072
(262)	(3)	(0.99969)	(0.99968)
[259]	[2]	[0.99967]	[0.99966]

5.7.3 Least Absolute Shrinkage and Selection Operator (LASSO) Model

The tolerance and variance inflection function showed the multicollinearity problem was presented in the regressors (Table 5.6). The aim of the present phases was to select the subset regression variables that variables could able to predict the final stature by using the LASSO model. The LASSO was a hybrid of a penalized estimation procedure and a variable selection procedure which not only helps to improve the prediction accuracy when dealing with multicollinearity data, but also carries several nice properties such as interpretability and numerical stability.

Firstly, the optimal LASSO parameter *s* was obtained by using the 10-fold cross-validated mean squared prediction error. Next, using the optimal value of *s*, the model was estimated and predicted with output its coefficient. Figure 5.6a showed the cross-validated error at each value of index and their corresponding standard error bands. From Figure 5.6a, one can saw that the speed of cross validity scores decreases dramatically until s = 0.1919. Thus, the optimum value of s was 0.1919. Since, the optimal fraction value was very small to one (i.e., s is 0.1919), the fitted model was not full model, which includes subset variables in the model.

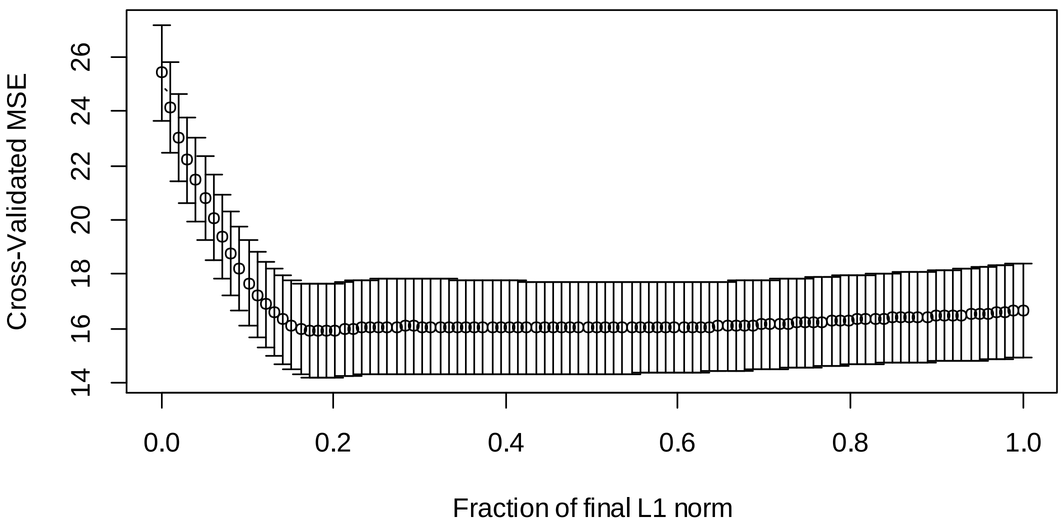


Figure 5.6a 10-fold cross-validated mean squared prediction error of boys

In order to saw how LASSO shrinks and predicts the coefficients more clearly, the LASSO estimated as a function of the standardized relative bound *s* as plotted. The *x*-axis (Figure 5.6b) on the graph presented as the ratio of the sum of the absolute current estimate over the sum of the absolute OLS estimates. And, for the *y*-axis (Figure 5.6b) was being standardized coefficients, generally when running LASSO, our *S* (data matrix of regressors) variables was standardized so that the penalization occurred equally over the variables. If they were measured on different scales, the penalization would be uneven. Intuitively, every coefficient would be squeezed to zero as *s* had gone to zero. Figure 5.6b showed that each monotone decreasing curve represented a coefficient as a function of relative bound *s*. The vertical red lines showed the optimum *s* value. The covariates enter the regression equation sequentially as *s* increased, in order i = 9, 14, 8, 13, 2, 12, 4, 6, 10, 1, 3, 7, 5, 11 and corresponding variables as S_{10}, *MS*, S_9, *FS*, S_3, S_{13}, S_5, S_7, S_{11}, S_2, S_4, S_8, S_6, S_{12}.

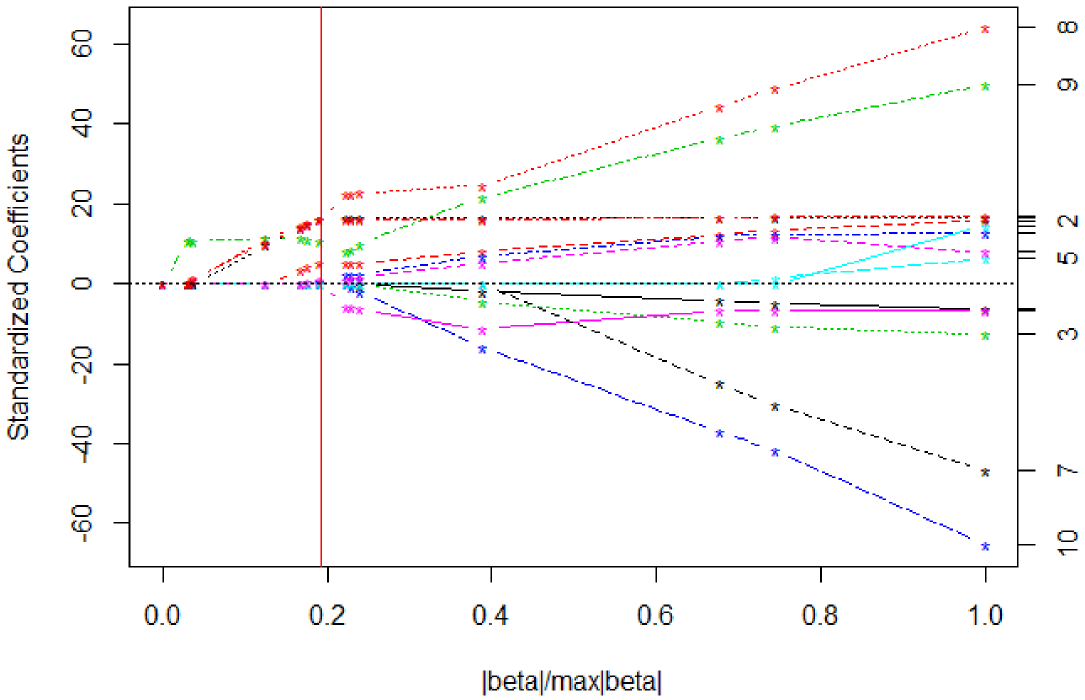


Figure 5.6b LASSO coefficient shrinkage of boys

Secondly, using the optimum value 0.1919, the LASSO model was estimated. The LASSO estimate, standard error, Z-score, probability of exact level of significant and comment of significant levels were shown in Table 5.10a. The standard errors (SE) were estimated by bootstrap resampling of residuals from the original data set. LASSO chooses S_3, S_5, S_7, S_9, S_{10}, S_{13}, *FS* and *MS*. Notice that LASSO yielded smaller standard error for *FS* and *MS*. This showed that LASSO predicted the coefficients with more accuracy. The *FS* and *MS* were found significant respective at less than 1%, imply that had impact on *PFS* for boys. Table 5.10a also showed a tendency that LASSO estimate was subset of all predictor. This was due to its constraint nature, that all predictions were subtracted by a threshold value.

Table 5.10a Summary statistics of LASSO model for boys

Predictor	Coefficients	Standard Error	Z-score	Pr(>|t|)	Significant at
S_2	0.0000	0.1065	0.0000	1.0000	
S_3	0.0839	0.1410	0.5946	0.5521	
S_4	0.0000	0.1924	0.0000	1.0000	
S_5	0.0076	0.2060	0.0370	0.9705	
S_6	0.0000	0.2029	0.0000	1.0000	
S_7	-0.0034	0.3387	-0.0102	0.9919	
S_8	0.0000	0.4534	0.0000	1.0000	
S_9	0.2345	0.4650	0.5043	0.6141	
S_{10}	0.1419	0.4002	0.3547	0.7228	
S_{11}	0.0000	0.3240	0.0000	1.0000	
S_{12}	0.0000	0.2807	0.0000	1.0000	
S_{13}	0.0121	0.1502	0.0802	0.9361	
FS	0.2284	0.0585	3.9060	0.0001	< 1 percent
MS	0.2724	0.0728	3.7434	0.0002	< 1 percent

Therefore, the prediction equations for boys using LASSO estimate as follows:

For boys:

$$PFS = 0.0839\, S_3 + 0.0076\, S_5 - 0.0034\, S_7 + 0.2345\, S_9 + 0.1419\, S_{10} + 0.0121\, S_{13} + 0.2284\, FS + 0.2724\, MS$$

In the above equation, only *FS* and *MS* were found statistically significant at less than 1% level.

Similarly, LASSO model was applied to the girl's data. From Figure 5.7a, one can saw that the speed of cross validity scores decreased dramatically until s = 0.7879. Thus, the value as 0.7879 was picked as an optimized s value. Since, the optimal fraction value was near to one (i.e., s is 0.7879), the fitted model was near to full model, which included subset variables in the model.

LASSO estimated as a function of the standardized relative bound s was plotted for seeing how LASSO shrinks and predicted the coefficients more clearly. Intuitively, every coefficient would be squeezed to zero as s goes to zero. Figure 5.7b showed that each monotone decreasing curve represented a coefficient as a function of relative bound s. The vertical red lines showed at which s value that each coefficient shrinks to zero. The covariates entered the regression equation sequentially as s increase, in order i = 12, 14, 13, 1, 10, 11, 8, 3, 2, 9, 5, 6, 4, 7 and corresponding variables as S_{13}, *MS*, *FS*, S_2, S_{11}, S_{12}, S_9, S_4, S_3, S_{10}, S_6, S_7, S_5, S_8.

Table 5.10b showed the LASSO estimate, standard error, Z-score and exact significant level of the parameters, using the optimum value of s was 0.7879. The standard errors (SE) were estimated by bootstrap resampling of residuals from the original data set. LASSO chooses S_2, S_3, S_4, S_5, S_6, S_7, S_{10}, S_{11}, S_{12}, S_{13}, *FS* and *MS*. Notice that, the LASSO yielded smaller standard error for *FS* and *MS*. This showed that LASSO predicted the coefficients with more accuracy. The S_{12}, S_{13} and *MS* were found significant respective at less than 1%, imply that having impact on *PFS* for girls. Table 5.10b also showed a near tendency that LASSO estimate was subset of all predictor. This was due to its constraint nature, that all predictions were subtracted by a threshold value.

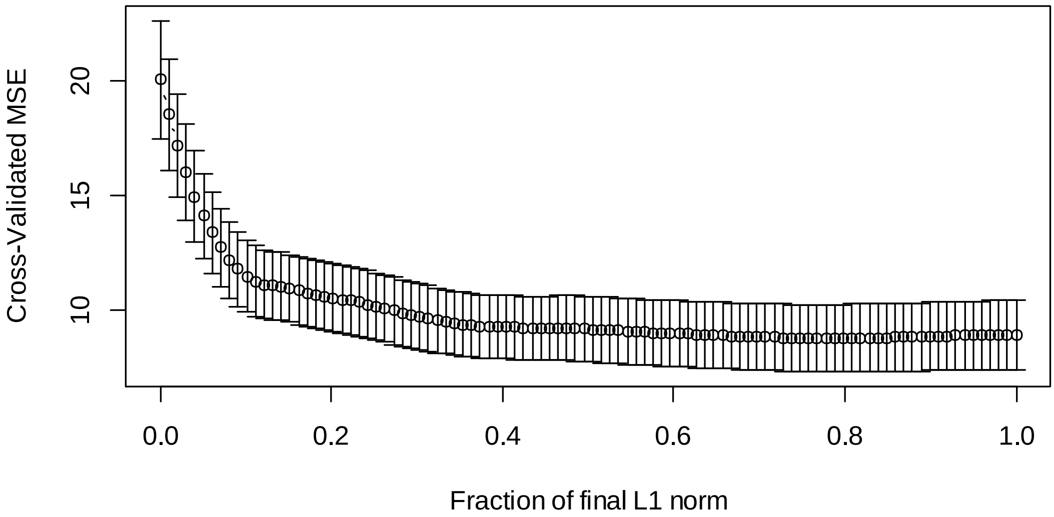


Figure 5.7a 10-fold cross-validated mean squared prediction errors of girls

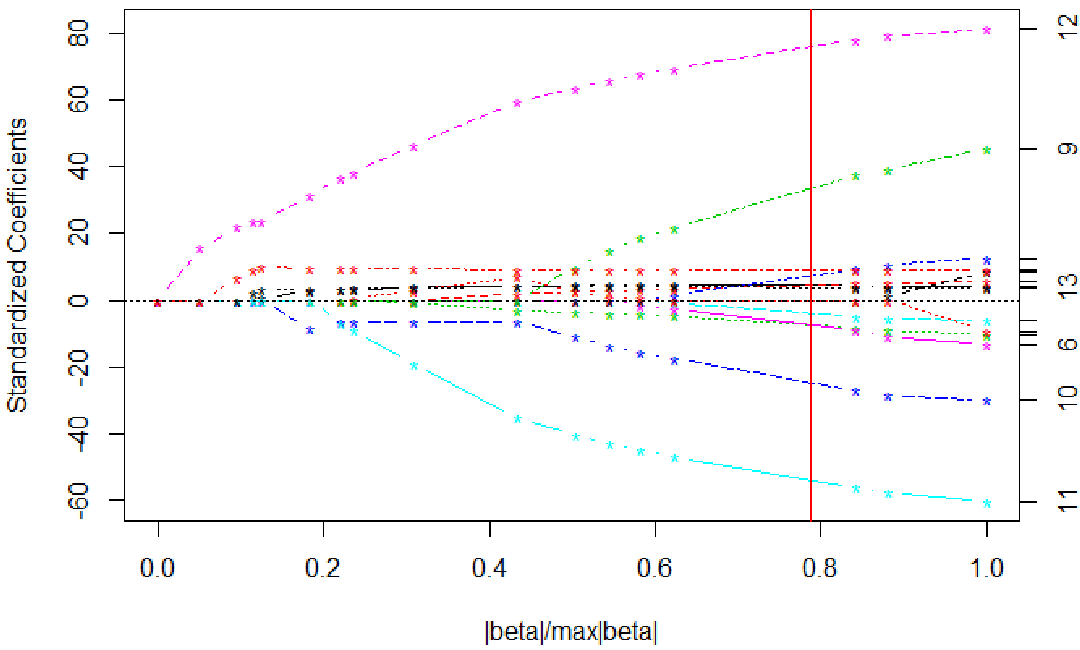


Figure 5.7b LASSO coefficient shrinkage of girls

Table 5.10b Summary statistics of LASSO model for girls

Predictor	Coefficients	Standard Error	Z-score	Pr(>|t|)	Significant at
S_2	0.1085	0.1186	0.9153	0.3601	
S_3	0.1036	0.1383	0.7490	0.4539	
S_4	-0.1468	0.1339	-1.0961	0.2730	
S_5	0.1377	0.1515	0.9087	0.3635	
S_6	-0.0656	0.1489	-0.4405	0.6596	
S_7	-0.1272	0.2569	-0.4952	0.6204	
S_8	0.0000	0.3690	0.0000	1.0000	
S_9	0.0000	0.3615	0.0000	1.0000	
S_{10}	0.5287	0.2899	1.8233	0.0683	
S_{11}	-0.3739	0.2355	-1.5878	0.1123	
S_{12}	-0.8443	0.2545	-3.3173	0.0009	< 1 percent
S_{13}	1.3490	0.1801	7.4907	0.0001	< 1 percent
FS	0.0750	0.0603	1.2439	0.2135	
MS	0.1579	0.0546	2.8934	0.0038	< 1 percent

Therefore, the prediction equations for girls using LASSO estimate as follows:

For girls:

$$PFS = 0.1085\,S_2 + 0.1036\,S_3 - 0.1468\,S_4 + 0.1377\,S_5 - 0.0656\,S_6 - 0.1272\,S_7 + 0.5287\,S_{10} - 0.3739\,S_{11}$$
$$- 0.8443\,S_{12} + 1.3490\,S_{13} + 0.0750\,FS + 0.1579\,MS$$

In the above equation, only S_{12}, S_{13} and *MS* were found statistically significant at less than 1% level.

Precision of the Estimated Model: Analysis of residual for boys and girls were considered to understand the precision of the predicted final stature. Average of observed, predicted value and residual values, 95% confidence bounds of mean residuals and standard error (SE) of predicting the final stature for boys and girls were shown in Table 5.11. The average value of residual was 0.00 cm and standard error of residuals and prediction were as 3.84 cm and 0.24 cm, respectively for boys that were small, implying the prediction equation was sufficient for predicting the final stature. Similarly, the average value of residual was 0.00 cm and standard error of residuals and prediction were as 2.59 cm and 0.32 cm, respectively for girls that were also small, implying the prediction equation was sufficient for predicting the final stature. Therefore, the prediction of final stature based on different stature variables using LASSO model accurately estimated the stature for both boys and girls.

Table 5.11 Average of observed, predicted, residual, 90% confidence bounds of residuals and standard error (SE) of predicted equations of final stature for boys and girls. Values in the first bracket ( ) were of Ali and Ohtsuki (2001) and that in third bracket [ ] were of Rahman and Ali (2003)

Data Set size	Observed stature (cm)	Predicted stature (cm)	Residual (cm)		95% confidence bound of mean residuals (cm)		SE of prediction (cm)
			Mean	SD	Lower	Upper	
			Boys				
180 (410) [464]	172.85 (172.34) [171.69]	172.85 (172.23) [171.58]	0.00 (0.113) [0.109]	3.84 (5.17) [4.80]	-0.57 (-6.024) [-0.258]	0.57 (7.497) [0.476]	0.24 (0.42) [0.30]
			Girls				
120 (262) [259]	158.96 (159.00) [158.44]	158.96 (159.19) [158.39]	0.00 (0.027) [0.046]	2.59 (2.80) [2.89]	-0.47 (-2.652) [-0.250]	0.47 (3.512) [0.344]	0.32 (0.29) [0.24]

5.7.4 Comparisons of Precision

Generally, the precision of least absolute shrinkage and selection operator (LASSO) model was better than that of the ridge regression model. The LASSO model produced less standard deviation of residuals and standard error of predicted stature than the forward stepwise ridge regression model. The forward stepwise ridge regression model under-estimated the *PFS* but LASSO accurately estimated the *PFS* for both boys and girls. Therefore, the LASSO fitted the model more adequately than the forward stepwise ridge regression model. The forward stepwise ridge regression model showed that the final prediction equations for boys and girls depended on 'stature at the age 9 and parent statures' and 'stature at the age 13 and parent statures', respectively. But, the LASSO model implied that the 'parent stature' for boys, and 'stature at the age 12, stature at the age 13 and mother stature' for girls had a significant effect on the predicted final stature (*PFS*). In LASSO model, some unnecessary (insignificance) variables were included in an estimated model but that of not do by the forward stepwise ridge regression model. Therefore, the equation of forward stepwise ridge regression model can be used for predicting final stature.

5.8 Comparisons with Other Studies

According to Jolicoeur et al. (1992), JPA-2 model fitted to the human stature data better than all other asymptotic models till 1991. While, BTT model was found to be better than JPA-2 model **(**Rahman et al., 2004). The present study showed that the proposed model was better than BTT model. Hence, it was argued that the proposed higher dimensional growth model was better than all other asymptotic models till now.

Comparing with Ali and Ohtsuki (2001), and Rahman and Ali (2003), the present study showed that the prediction of final stature based on statures at different ages and parent statures were more useful with higher precession for the Japanese population. Statures at the ages 9 and 13 were very important for *PFS*, respectively for boys and girls which was in accord with others (Ali and Ohtsuki, 2001; and Ali and Rahman, 2003). Here, the parent statures were also important for *PFS* for both boys and girls. On the other hand, the present prediction equations were easy to calculate and need three stature values only and need not to fit any model with longitudinal data from birth to maturity.

The mean residuals of the predicted final stature of the present study (Table 5.8 and Table 5.11) were smaller compared with that of others (Bayley and Pinneau, 1952; Khamis and Guo, 1993; Khamis and Roche, 1994; Roche et al., 1975a, b; Wainer et al., 1978; Ali and Ohtsuki, 2001; Rahman and Ali, 2003). Moreover, the LASSO estimated *PFS* more precisely than the forward stepwise ridge regression model. Average prediction failure was reported by Khamis and Guo (1993) as about 10% for boys and 8% for girls. Ali and Ohtsuki (2001) found 12% prediction failure for case 3 (who do not have a mid-growth spurt) of Japanese boys and 7% for case 1 (whole sample individuals), 2% for case 2 (who have the mid-growth spurt), and 6% for case 3 (who do not have the mid-growth spurt) of Japanese girls. While, Rahman and Ali (2003) found in boys, 8% failure occurred for case 3 (who do not have a mid-growth spurt), but, in girls, failures of 6% for case 1 (whole sample individuals), 2% for case 2 (who have the mid-growth spurt), 4% for case 3 (who do not have the mid-growth spurt). In the present prediction, the average prediction failure was 0% for both boys and girls.

Standard errors of the predicting final stature (Table 5.8 and Table 5.11) were smaller than those of some others (Onat, 1975 and 1983; Ali and Ohtsuki, 2001; Rahman and Ali, 2003). Comparing with 90% confidence bounds for residuals, the girls (Table 5.8 and Table 5.11) showed better prediction than those of some others (Ali and Ohtsuki, 2001; Rahman and Ali, 2003; Khamis and Roche, 1994; Roche et al., 1975a; Wainer et al., 1978).

Finally, the values of R^2 and cross validity predictive power were very high for both boys and girls compared with Ali and Ohtsuki (2001), and Rahman and Ali (2003). Therefore, the proposed prediction equations are also useful for predicting final stature of Japanese population.

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

"Give your decisions, never your reasons; your decisions may be right, but your reasons are sure to be wrong." - Willian Murray

CHAPTER 6

CONCLUSION AND RECOMMENDATION

6.1 Outline

This chapter includes major finding with recommendation, limitation of study and scope of further study.

6.2 Major Finding

The secondary longitudinal data of age, weight and stature of 300 Japanese (180 boys and 120 girls), each between 0 to 20 years old and covering birth-years of 1967 to 1977, have been used. The estimated population mean, standard deviation (SD), and correlation matrix of the parameters of proposed model and their average root mean square errors for boys and girls were shown in Result and Discussion section. In the present study, our proposed model demonstrated that, on average, 30.102%, 29.933%, and 39.965% of the total final stature were completed during early, middle and adolescent phase of growth, respectively, for the male population and for the female population, these percentages are 29.169%, 36.137%, and 34.694%, respectively. The correlation between (a_1, a_2), (a_1, a_{21}), (a_1, a_{22}), (a_1, c_2), (a_{11}, a_{12}), (a_{11}, c_1), (a_{12}, c_1), (a_{21}, a_{22}), (a_{21}, c_2), (a_{22}, c_2), (a_{31}, a_{32}), (a_{31}, c_3) and (a_{32}, c_3) for boys, and that of between (a_1, a_2), (a_1, a_3), (a_{11}, a_{12}), (a_{11}, c_1), (a_{12}, c_1), (a_{12}, a_3), (a_3,

c_1), (a_2, a_{22}), (a_2, c_2), (a_{21}, c_2), (a_3, c_3), (a_{31}, a_{32}) and (a_{32}, c_3) for girls are statistically significant at most 5%. The average root mean square error for boys is larger than that for girls for both boys and girls. The average root mean square error of BTT model is larger than that of the proposed model for both boys and girls.

The distribution of predicted stature, one average, showed that the boys became taller than girls from age 1 to 9 and 12 to 25. But, the distribution of predicted stature, one average, showed that the girls became taller than boys from age 10 to 11.

The line graph of observed *vs.* fitted showed that the proposed model estimated the stature for boys and girls more precisely than the BTT model. A particular data set, the mean square error of the BTT model is always greater than the mean square error of the proposed model for different values of shape parameters, implying that the proposed model estimated the stature more precisely than the BTT model for different values of shape parameters also. For the different data set, the value of shape parameters assumed to be fixed and thus it is found that the mean square error of the proposed model were always smaller than the BTT model for both boys and girls. Therefore, the proposed model estimated the stature for both boys and girls more precisely than the BTT model.

After extracting the final stature and stature at different ages from the well fitted proposed model and eliminating the problem of multicollinearity using Forward

Stepwise Ridge Regression and LASSO techniques, the following equations were established with higher precession, validity and stability:

From Forward Stepwise Ridge Regression Technique

For boys: $PFS = 0.333572\ MS + 0.342368\ S_9 + 0.322908\ FS$

For girls: $PFS = 0.387948\ S_{13} + 0.316876\ MS + 0.294889\ FS$

From Least Absolute Shrinkage and Selection Operator (LASSO) Technique

For boys:

$$PFS = 0.0839\, S_3 + 0.0076\, S_5 - 0.0034\, S_7 + 0.2345\, S_9 + 0.1419\, S_{10} + 0.0121\, S_{13} + 0.2284\, FS + 0.2724\, MS$$

In the above equation, only *FS* and *MS* were found significant at less than 1% level.

For girls:

$$PFS = 0.1085\, S_2 + 0.1036\, S_3 - 0.1468\, S_4 + 0.1377\, S_5 - 0.0656\, S_6 - 0.1272\, S_7 + 0.5287\, S_{10} - 0.3739\, S_{11}$$
$$- 0.8443\, S_{12} + 1.3490\, S_{13} + 0.0750\, FS + 0.1579\, MS$$

In the above equation, only S_{12}, S_{13} and *MS* were found significant at less than 1% level.

The above equations need not any longitudinal data, but also stature at age 9, 12, 13 and parent statures to estimate final stature.

6.3 Limitations of the Study

In the present study, only two regressor variables, i.e., age and weight have been considered and many other important variables on stature have been omitted. It would have been better if we could have included all the variables in the model, thereby better explaining the stature. Due to the limitation of research title and lack of data, only three variables have been considered. Finally, in age and weight there may be multicollinearity problems which are ignored.

6.4 Scope for Further Research

In further research, one can try to find out the asymptotic distribution of the growth parameters of the proposed model. Longitudinal data of human growth have not been collected and preserved in our country. After collecting this type of data our country may also predict the final stature of our children. Moreover, the present study considers three variables, and the prediction is based on the ages 2 to 13 and parent statures. In future, there’s still plenty of scope for including more than three variables in the model, and the prediction of final stature may also be carried out, based on the biological parameters in the days to come.

BIBLIOGRAPHY

"A Mathematician knows how to solve a problem but he can't do it." - W.E. Milne

BIBLIOGRAPHY

Ali MA and Ohtsuki F (2000). Estimation of maximum increment age in height and weight during adolescence and the effect of World War II. *American Journal of Human Biology*, 12:363–370.

Ali MA and Ohtsuki F (2001). Prediction of adult stature for Japanese population: A stepwise regression approach. *American Journal of Human Biology*, 13:316–322.

Ali MA, Rahman JAMS, Ashizawa K and Ohtsuki F (2004a). Human growth in Japanese children: An application of triphasic generalized logistic model. *International Journal of Statistical Sciences*, 3:75–92.

Ali MA, Rahman JAMS, Ashizawa K and Ohtsuki F (2004b). Stepwise regression for predicting final stature of Japanese children. *International Journal of Statistical Sciences*, 3:269–280.

Ashizawa K, Kumakura C, and Kusumoto A (1998). Growth of the Philippine children in Reference to socioeconomic environment. *Anthropological Science*, 106:77-94.

Ashizawa K, Kato S and Eto M (1993). Individual adolescent growth of stature, body weight, and chest circumference of girls in Tokyo. *Anthropological Science*, 102(4):421–446.

Ashizawa K, Kumakura C, Kato S, Kawabe T, Hauspie RC and Eto M (2005). Adolescent height growth of girls in Tokyo. *Anthropological Science*, 113:245–252.

Bayley N and Pinneau SR (1952). Tables for predicting adult height from skeletal age: Revised for use with Greulich-Pyle hand standards. *Journal of Pediatrics*, *40*:423-441.

Berkey CS (1982). Comparison of two longitudinal growth models for preschool children. *Biometrics*, 38:221–234.

Berkey CS and Reed RB (1987). A model for describing normal and abnormal growth in early childhood. *Human Biology*, 49:973–987.

Billewicz WZ and McGregor IA (1982). A birth-to-maturity longitudinal study of heights and weights in two West African (Gambian) villages, 1951-1975. *Annals of Human Biology*. 9:309-320.

Bjelica D, Popovi S, Kezunovi M, Petkovi J, Jurak G and Grasgruber P (2012). Body height and its estimation utilizing arm span measurements in Montenegrin adults. *Anthropological Notebooks*, 18(2):69–83.

Bock RD, Wainer H, Petersen A, Thissen D, Murray J and Roche AF (1973). A parameterization for individual human growth curves. *Human Biology*, 45:63–80.

Bock RD and Thissen D (1976). Fitting multi-component models for growth in stature. *Proceeding of the 9th International Biometrics Conference*, 1:431–442.

Bock RD and Thissen D (1980). Statistical problem of fitting individual growth curves. In: Johnston FE, Roche AF, Susanne C (eds.), *Human physical growth and maturation: methodologies and factors,* New York, 30:265–290.

Bock RD, du Toit SHC and Thissen D (1994). AUXAL: *Auxological analysis of longitudinal measurements of human stature*. Chicago: Scientific software international.

Bogin B and Silva MIV (2010). Leg length, body proportion, and health a review with a note on beauty. *International Journal of Environmental Research and Public Health*, 7:1047–1075. doi:10.3390/ijerph7031047

Bogin B, Wall M and MacVean RB (1990). Longitudinal growth of high socioeconomic status Guatemalan children analized by the Preece-Baines function: An international comparison. *Annals of Human Biology*, *2*:271-281.

Bogin B, Wall M and MacVean RB (1992). Longitudinal analysis of adolescent growth of *L*adino and Mayan school children in Guatemala: Effects of environment and sex. *American Journal of Physical Anthropology*, *89*:447-457.

Brown T and Townsend GC (1982). Adolescent growth in height of Australian Aboriginals analysed by the Preece-Baines function: a longitudinal study. *Annals of Human Biology*, *9*:495-505.

Byard PJ, Guo S and Roche AF (1993). Family resemblance for Preece-Baines growth curve parameters in the Fels Longitudinal Growth Study. *American Journal of Human Biology*, *5*:151-157.

Box GEP and Cox DR (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, Series B (Methodological), 26:211–243.

Broeck JVD, Massa GG, Attanasio A, Matranga A, Chaussain JL, Price DA, Aarskog D, Wit JM and European Study Group (1995). Final height after long-term growth hormone treatment in turner syndrome. *The Journal of Pediatrics*, 127:729–35.

Cameron N, Tanner JM and Whitehouse RH (1982). A longitudinal analysis of the growth of limb segments in adolescence. *Annals of Human Biology, 9*:211-220.

Chittawatanarat K, Pruenglampoo S, Trakulhoon V, Ungpinitpong W and Patumanond J (2012). Height prediction from anthropometric length parameters in Thai people. *Asia Pacific Journal of Clinical Nutrition*, 21 (3):347–354.

Choksi A, Babu GR and Dahiya MS (2014). Study on the primary characteristics of identification estimation of stature from palm length among the native Guajarati population. *NHL Journal of Medical Sciences*, 3(2):13–17.

Chumlea WC, Guo SS, Wholihan K, Cockram D, Kuczmarski RJ and Johnson CL (1998). Stature prediction equations for elderly non-Hispanic white, non-Hispanic black, and Mexican-American persons developed from NHANES III data. *Journal of the American Dietetic Association*, 98(2):137–142.

Cole TJ (2000). Galton's midparent height revisited. *Annals of Human Biology*, 27(4):401–405.

CookRD (1977). Detection of influential observations in linear regression. *Technometrics*, 19(1):15–18. doi:10.2307/1268249.

Cook RD and Weisberg S (1982). *Residuals and influence in regression*, New York: Chapman and Hall.

Count EW (1943). Growth patterns of human physique: an approach to kinetic anthropometry. *Human Biology*, 15:1–32.

Csukás A, Takai S and Baran S (2006). Adolescent growth in main somatometric traits of Japanese boys: Ogi longitudinal growth study. *HOMO—Journal of Comparative Human Biology*, 57:73–86.

Deming J (1957). Application of the Gompertz curve to the observed pattern of growth in length of 48 individual boys and girls during the adolescent cycle of growth. *Human Biology*, 29:83–122.

Deming J, and Washburn AH (1963). Application of the Jenss curve to the observed pattern of growth during the first eight years of life in forty boys and forty girls. *Human Biology*, 35:484–506.

du Toit (1992). *Multi-component models for non-consecutive data with ARMA(1,1) residuals*. Department of Statistics, University of Pretoria.

Feldesman MR (1992). Femur stature ratio and estimates of stature in children. *American Journal of Physical Anthropology*, 87:447–459.

Fukami M, Matsuo N, Hasegawa T, Sato S and Ogata T (2003). Longitudinal auxological study in a female with SHOX (short stature homeobox containing gene) haploinsufficiency and normal ovarian function. *European Journal of Endocrinology*, 149:337–341.

Greulich WW and Pyle SI (1950). *Atlas of Skeletal Development of the Hand and Wrist*. Stanford University Press.

Hauspie RC (1980). Adolescent growth. In Johnston FE, Roche AF, and Susanne C (eds.): *Human Physical Growth and Maturation*. New York: Plenum, 161-175.

Hauspie RC, Das SR, Preece MA, and Tanner JM (1980a). A longitudinal study of the growth in height of boys and girls of West Bengal (India) aged six months to 20 years. *Annals of Human Biology*, 7:429-441.

Hauspie RC, Wachholder A, Baron G, Cantraine F, Susanne C and Graffar M (1980b). A comparative study of the fit of four different functions to longitudinal data of growth in height of Belgian girls. *Annals of Human Biology*, 7:347-358.

Henneberg GM and Louw JG (1998). Cross sectional survey of growth of urban and rural "cape coloured" school children: Anthropometry and functional tests. *American Journal of Human Biology*, 10:73–85.

Hoerl AE, Kennard RW and Baldwin KF (1975). Ridge regression: some simulation. *Communications in Statistics*, 4:105–123.

Hui L, XinNan Z, Jing Z, and ZongHan Z (2011). Physical growth of children in urban, suburban and rural mainland china a study of 20 years' change. *Biomedical and Environmental Science*, 24(1):1–11.

Huxley JS (1932). *Problems of relative growth*. Lincoln Mac Veagh—The Dial Press, New York.

Huxley JS and Tessier G (1936). Terminology of relative growth. *Nature*, 137:780–781.

Ilayperuma I, Nanayakkara G and Palahepitiya N (2010). A Model for the estimation of personal stature from the length of forearm. *International Journal of Morphology*, 28(4):1081–1086.

Israelsohn WJ (1960). Description and modes of analysis of human growth. *In Human Growth*, J.M. Tanner (ed.), 21-42. New York: Pergamon Press.

Jenss RM, and Bayley N (1937). A mathematical model for studying the growth of a child. *Human Biology*, 9:556–563.

Johnson ML (1993). Methods for the analysis of saltation and stasis in human growth data. In: Lampl M (ed). *Saltation and Stasis in Human Growth and Development: Evidence, Methods and Theory*. London: Smith-Gordon, 1999:27–32.

Johnson W, Choh AC, Soloway LE, Czerwinski SA, Towne B and Demerath EW (2011). Eighty-year trends in infant weight and length growth: The Felslongitudinal study. *The Journal of Pediatrics*, doi: 10.1016/j.jpeds.2011.11.002.

Jolicoeur P, Pontier J, Pernin MO and Sempe M (1988). A lifetime asymptotic growth curve for human height. *Biometrics*, 44:995–1003.

Jolicoeur P, Pontier J, and Abidi H (1992). Asymptotic models for the longitudinal growth of human stature. *American Journal of Human Biology*, 4:461–468.

Jones ADGB, Eisenmann JC and Sherar LB (2005). Controlling for maturation in pediatric exercise science. *Pediatric Exercise Science*, 17:18–30.

Kanefuji K and Shohoji T (1990). On a growth model of human height. *Growth development and Aging*, 54:155–165.

Karlberg J (1989). A biologically-oriented mathematical model (ICP) for human growth. *Acta Paediatrica Supplement*, 350:70–94.

Kato S, Ashizawa K and Satoh K (1998). An examination of the definition 'final height' for practical use. *Annals of Human Biology*, 25(3):263–70.

Khamis HJ and Guo S (1993). Improvement in the Roche-Wainer-Thissen stature prediction model: A comparative study. *American Journal of Human Biology*, 5: 669-679.

Khamis HJ and Roche AF (1994). Prediction adult stature without using skeletal age: The Khamis Roche method. *Pediatrics* 94(4):504–507.

Khamis HJ (1993). Improvement in the Roche-Wainer-Thissen stature prediction model: A comparative study. *American Journal of Human Biology*, 5(6):669–679.

Kuninaka H and Matsushita M (2014). Multiplicative modeling of children’s growth and its statistical properties. *Journal of the Physical Society of Japan*. Draft

Ledford AW and Cole TJ (1998). Mathematical models of growth in stature throughout childhood. *Annals of Human Biology*, 25:101–115.

Leigh SR (2001). Evolution of human growth evolutionary. *Anthropology*, 10:223–236.

Lindgren WG and Hauspie CR (1989). Heights and weights of Swedish school children born in 1955 and 1967. *Journal of Human Biology*, 16(5):397–406.

Maijanen and Niskanen (2006). Comparing stature-estimation methods on medieval inhabitants of Westerhus, Sweden. *Fennoscandia archaeologica*, 23:37–46.

Manwani AH and Agarwal KN (1973). The growth patterns of Indian infants during the first year of life. *Human Biology*, 45:341–349.

Marquardt DW (1970). Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics*, 12:591–612.

Marubini E, Resele LF and Barghini G (1971). A comparative fitting of the Gompertz and logistic functions to longitudinal height data during adolescence in girls. *Human Biology*, 43:237–252.

Mata LJ (1978). *The children of Santa Maria Cauque*. Cambridge, Massachusetts: MIT Press.

Merrell M (1931). The relationship of individual growth to average growth. *Human Biology*, 3:37–70.

Mirwald RL, Bailey DA, Cameron N and Rasmussen RL (1981). Longitudinal comparison of aerobic power in active and inactive boys aged 7.0 to 17.0 years. *Annals of Human Biology*, 8:405-414.

Mounir GM, Sayed NAE, Mahdy NH and Khamis SE (2007). Nutritional factors affecting the menarcheal state of adolescent school girls in Alexandria. *Journal of the Egyptian Public Health Association*, 82:239–260.

Needlman RD (2003). *Growth and development*. In: Behrman RE, et al. Nelson Textbook of Pediatrics. 17th ed. Philadelphia, Pa.: Saunders.

Nocedal J and Wright SJ (1999). *Numerical optimization*. Springer-Verlag. ISBN 0-387-98793-2.

Onat T (1975). Prediction of adult height of girls based on the percentage of adult height at onset of secondary sexual characteristics, at chronological age, and skeletal age. *Human Biology*, 47(1):117–130.

Onat T (1983). Multifactorial prediction of adult height of girls during early adolescence allowing for genetic potential, skeletal and sexual maturity. *Human Biology*, 55(2):443–461.

Pomeroy E and Stock JT (2012). Estimation of stature and body mass from the skeleton among coastal and mid-altitude Andean populations. *American Journal of Physical Anthropology*, 147:264–279.

Preece MA and Baines MJ (1978). A new family of mathematical models describing the human growth curve. *Annals of Human Biology*, 5:1–24.

Qin T, Shohoji T and Sumiya T (1996). Relationship between adult stature and timing of the pubertal growth spurt. *American Journal of Human Biology*, *8*:417-426.

Rahmandad (2014). Human growth and body weight dynamics: An integrative systems model. *PLoS ONE* 9(12): e114609. doi:10.1371/journal.pone.0114609

Raxter MH, Auerbach BM and Ruff CB (2006). Revision of the fully technique for estimating statures. *American Journal of Physical Anthropology*,130:374–384.

Rahman and Ali (2003). Prediction of adult stature of Japanese boys and girls. *Unpublished Ph.D. Dissertation*. Department of Statistics, University of Rajshahi, Rajshahi-6205, Bangladesh.

Rahman JAMS, Ali MA, Ashizawa K and Ohtsuki F (2004). Prediction of adult stature for Japanese population: an improvement of Ali-Ohtsuki equations. *Anthropological science*, 112:61–66.

Robertson TB (1908). On the normal rate of growth of an individual, and its biochemical significance. *Archiv fur Entwicklungsmechanik der Organismen*, 25:581–614.

Roche AF, Wainer H and Thissen D (1975a). The RWT method for the prediction of adult stature. *Pediatrics*, 56(6):1027–1033.

Roche AF, Wainer H, and Thissen D (1975b). Predicting adult stature for individuals. *Monograph in Paediatrics*. *3*: Karger, Basel.

Rosique J and Rebato E (1995). Comparative study of statural growth in Spanish populations. *American Journal of Human Biology*, 7:553–564.

Sarajlić N, Cihlarž Z, Klonowski EE and Selak I (2006). Stature estimation for Bosniam male population. *Bosnian Journal of Basic Medical Sciences*, 6(1):62–67.

Shahar S and Pooy NS (2003). Predictive equations for estimation of stature in Malaysian eldered people. *Asia Pacific Journal Clinical Nutrition*, 12(1):80–84

Shahin MA, Ali MA, Ali ABMS (2012). Higher Dimensional Growth Model: An Extension of the Bock, Thissen and du Toit (BTT) Model. *Proceedings of International Conference on Statistical Data Mining for Bioinformatics, Health, Agriculture and Environment.* 21-24 December, 2012, 301-309, ISBN: 978-984-33-5876-9.

Shahin MA, Ali MA and Ali ABMS (2013). An extension of generalized triphasic logistic human growth model. *Journal of Biometrics and Biostatistics*, 4:162. doi:10.4172/2155- 6180.1000162

Shohoji T and Sasaki H (1987). Individual growth of Japanese. *Growth*, 51:432–450.

Sorensen D, Gianola D (2002). *Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics*, Springer-Verlag, New York.

Stovitz SD, Pereira MA, Vazquez G, Lytle LA and Himes JH (2008). The interaction of childhood height and childhood BMI in the prediction of young adult BMI. *Obesity*, 16:2336–2341. doi:10.1038/oby.2008.359

Stevens J (1996). *Applied multivariate statistics for the social sciences*. Third Edition. Lawrence Erlbaum Associates, Inc., Publishers, New Jersey.

Sunil, Dikshit PC and Aggrawal A (2005). Estimation of stature from hand length. *Journal of Indian Academy of Forensic Medicine*, 27(4):219–221.

Suski PM and Angeles L (1935). Body measurements of Japanese children born in America. *California and Western Medicine*, 43(3):208–209.

Takai S (1993). Velocities for the Tanner-Whitehouse 2 skeletal maturity in northwest Japanese children. *Okajimas Folia Anatomica Japonica,* 70(2-3):119–126.

Tanner JM, Hayashi T, Preece MA and Cameron N (1982). Increase in length of leg relative to trunk in Japanese children and adults from 1957 to 1977: comparison with British and with Japanese Americans. *Annals of Human Biology,* 9:411-423.

Tanner JM, Healy MJR, Lockhart RD, MacKenzie JD, and Whitehouse RH (1956). The prediction of adult body measurements from measurements taken each year from birth to five years. *Archives of Disease in Childhood*, 31:372-381.

Tibshirani R (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society*, Series B (Methodological), 58(1):267–288.

Tsuzaki S, Matsuo N, Saito M and Osano M (1990). The head circumference growth curve for Japanese children between 0-4 years of age: Comparision with Caucasian children and correlation with stature. *Annals of Human Biology*, 17(4):297–303.

Wainer H, Roche AF and Bell S (1978). Predicting adult stature without skeletal age and without paternal data. *Pediatrics*, 61(4):569–572.

Wingerd J (1970). The relation of growth from birth to two years to sex, parental size and other factors, using Rao's method of transformed time scale. *Human Biology*, 42: 105-131.

Zemel B and Johnston F (1994). Application of the Preece-Baines growth model to cross-sectional data: Problems of validity and interpretation. *American Journal of Human Biology*, 6:563-570.

APPENDIX-1

APPENDIX-1

SAMPLE DATA SET

All the data sets (total data set number 300) are available to the author. But, only one data set for the Japanese boys and one for girls are shown bellow.

Table: Age, stature, weight, father stature and mother stature of Japanese boys and girls

Boys					Girls				
Data Set Number 1					Data Set Number 1				
Age (year)	**Stature (cm)**	**Weight (kg)**	**Father Stature (cm)**	**Mother Stature (cm)**	**Age (year)**	**Stature (cm)**	**Weight (kg)**	**Father Stature (cm)**	**Mother Stature (cm)**
01	75.0	10.0	171.0	155.5	01	70.6	8.6	168.0	156.0
02	83.0	11.7			02	80.6	10.1		
03	91.8	13.7			03	92.2	13.5		
04	95.8	14.3			04	96.0	14.5		
05	102.8	16.2			05	99.2	15.4		
06	107.3	17.4			06	106.0	16.6		
07	116.7	20.0			07	110.0	18.5		
08	123.5	21.5			08	115.4	20.5		
09	126.9	24.0			09	120.6	22.4		
10	134.6	25.5			10	125.8	25.1		
11	141.3	27.8			11	131.1	26.8		
12	145.3	30.0			12	137.4	29.6		
13	153.5	37.5			13	141.1	32.5		
14	161.3	43.0			14	149.5	39.4		
15	166.3	45.5			15	153.8	41.2		
16	168.0	47.0			16	155.4	44.4		
17	169.0	50.0			17	156.1	43.0		
18	169.9	51.0			18	157.1	44.0		
19	169.8	53.0			19	157.6	45.0		
20	170.5	56.0			20	---	---		

APPENDIX-2

APPENDIX-2

MEAN SQUARE ERRORS

Table 1 Mean square errors for BTT and proposed models for boys

D.S. No.	MSE BTT Model	MSE Proposed Model	D.S. No.	MSE BTT Model	MSE Proposed Model	D.S. No.	MSE BTT Model	MSE Proposed Model
01	0.81	0.46	**61**	0.89	0.82	**121**	0.73	0.41
02	0.49	0.42	**62**	0.72	0.28	**122**	0.46	0.40
03	0.47	0.34	**63**	0.78	0.65	**123**	0.73	0.34
04	0.63	0.19	**64**	0.41	0.27	**124**	0.30	0.15
05	0.81	0.72	**65**	0.56	0.32	**125**	0.25	0.18
06	0.46	0.34	**66**	0.40	0.33	**126**	1.74	0.79
07	0.85	0.60	**67**	1.38	0.34	**127**	0.30	0.28
08	0.42	0.33	**68**	0.68	0.24	**128**	0.77	0.66
09	0.89	0.31	**69**	0.65	0.62	**129**	0.76	0.38
10	0.49	0.36	**70**	0.85	0.59	**130**	0.98	0.92
11	0.74	0.70	**71**	0.48	0.41	**131**	0.67	0.35
12	0.88	0.67	**72**	0.64	0.27	**132**	0.59	0.25
13	0.97	0.92	**73**	0.58	0.18	**133**	0.99	0.64
14	1.00	0.75	**74**	0.45	0.35	**134**	0.62	0.39
15	0.56	0.48	**75**	0.34	0.31	**135**	0.58	0.25
16	1.41	0.51	**76**	0.62	0.38	**136**	0.45	0.35
17	0.96	0.72	**77**	1.02	0.66	**137**	0.34	0.31
18	1.14	0.97	**78**	0.60	0.53	**138**	1.00	0.42
19	0.29	0.28	**79**	0.62	0.32	**139**	0.58	0.42
20	0.75	0.43	**80**	0.67	0.36	**140**	1.14	0.92
21	0.56	0.17	**81**	1.38	0.42	**141**	0.60	0.53
22	0.59	0.26	**82**	0.75	0.38	**142**	1.02	0.93
23	0.87	0.51	**83**	0.77	0.66	**143**	1.30	1.11
24	0.92	0.51	**84**	0.30	0.28	**144**	0.49	0.44
25	0.84	0.34	**85**	0.60	0.40	**145**	0.44	0.31
26	0.83	0.50	**86**	0.44	0.39	**146**	0.62	0.47
27	0.77	0.33	**87**	0.62	0.47	**147**	0.65	0.62
28	0.74	0.45	**88**	0.60	0.53	**148**	1.06	0.88
29	1.33	0.48	**89**	0.49	0.30	**149**	0.83	0.81
30	0.97	0.53	**90**	0.81	0.71	**150**	0.51	0.38
31	0.78	0.36	**91**	0.67	0.29	**151**	0.52	0.46
32	0.63	0.29	**92**	0.87	0.65	**152**	0.59	0.42
33	0.87	0.33	**93**	0.35	0.19	**153**	0.70	0.67
34	0.44	0.31	**94**	0.44	0.35	**154**	0.35	0.15
35	1.02	0.93	**95**	0.77	0.57	**155**	0.71	0.52
36	0.48	0.40	**96**	0.47	0.45	**156**	0.82	0.67

Continued…

D.S. No.	MSE		D.S. No.	MSE		D.S. No.	MSE	
	BTT Model	Proposed Model		BTT Model	Proposed Model		BTT Model	Proposed Model
37	0.82	0.47	**97**	0.75	0.33	**157**	0.39	0.22
38	0.65	0.59	**98**	0.56	0.30	**158**	0.54	0.30
39	0.72	0.53	**99**	0.74	0.32	**159**	0.25	0.19
40	0.34	0.27	**100**	0.41	0.36	**160**	0.69	0.35
41	0.73	0.64	**101**	0.49	0.30	**161**	0.48	0.28
42	0.64	0.51	**102**	0.73	0.46	**162**	1.04	0.91
43	0.57	0.22	**103**	0.51	0.31	**163**	0.48	0.34
44	1.06	0.91	**104**	1.30	0.83	**164**	0.53	0.28
45	0.92	0.79	**105**	0.63	0.60	**165**	0.77	0.29
46	0.52	0.33	**106**	0.69	0.60	**166**	0.93	0.83
47	0.52	0.43	**107**	0.82	0.71	**167**	0.97	0.61
48	0.74	0.30	**108**	0.35	0.22	**168**	0.73	0.48
49	0.71	0.30	**109**	0.31	0.23	**169**	0.63	0.33
50	0.60	0.44	**110**	0.82	0.72	**170**	0.39	0.26
51	0.89	0.77	**111**	0.51	0.43	**171**	0.89	0.76
52	0.82	0.31	**112**	0.74	0.41	**172**	0.40	0.11
53	0.39	0.25	**113**	0.68	0.41	**173**	0.66	0.30
54	0.73	0.26	**114**	0.75	0.32	**174**	1.40	0.34
55	0.66	0.56	**115**	0.42	0.30	**175**	0.40	0.32
56	0.94	0.60	**116**	0.75	0.33	**176**	0.41	0.30
57	0.53	0.28	**117**	0.94	0.53	**177**	0.41	0.27
58	0.77	0.29	**118**	0.41	0.26	**178**	0.56	0.41
59	0.40	0.20	**119**	0.74	0.33	**179**	0.40	0.34
60	0.36	0.30	**120**	0.67	0.50	**180**	0.79	0.31

Note: D.S. No. means Data Set Number

Table 2 Mean square errors for BTT and proposed models for girls

D.S. No.	MSE		D.S. No.	MSE		D.S. No.	MSE	
	BTT Model	Proposed Model		BTT Model	Proposed Model		BTT Model	Proposed Model
01	0.70	0.48	**41**	0.29	0.16	**81**	0.62	0.43
02	0.38	0.36	**42**	0.53	0.38	**82**	0.77	0.47
03	0.46	0.44	**43**	0.56	0.43	**83**	1.06	0.68
04	0.69	0.44	**44**	0.46	0.31	**84**	1.18	0.88
05	0.40	0.39	**45**	0.69	0.59	**85**	0.57	0.41
06	0.43	0.32	**46**	0.46	0.41	**86**	0.51	0.36
07	0.45	0.24	**47**	0.29	0.25	**87**	0.42	0.25
08	0.28	0.24	**48**	0.50	0.34	**88**	0.30	0.22
09	0.49	0.28	**49**	0.59	0.38	**89**	1.06	0.31
10	0.73	0.49	**50**	0.68	0.32	**90**	0.89	0.63
11	0.26	0.21	**51**	0.43	0.40	**91**	0.36	0.28
12	0.46	0.24	**52**	0.47	0.21	**92**	0.31	0.26
13	0.74	0.44	**53**	0.77	0.48	**93**	0.71	0.41
14	0.31	0.19	**54**	0.71	0.57	**94**	0.81	0.56
15	0.79	0.58	**55**	0.26	0.21	**95**	0.32	0.19
16	1.03	0.30	**56**	0.88	0.21	**96**	1.07	0.75
17	0.76	0.58	**57**	0.38	0.22	**97**	0.61	0.42
18	0.31	0.19	**58**	0.46	0.39	**98**	0.78	0.57

Continued…

D.S. No.	MSE		D.S. No.	MSE		D.S. No.	MSE	
	BTT Model	Proposed Model		BTT Model	Proposed Model		BTT Model	Proposed Model
19	1.14	0.30	**59**	1.05	0.40	**99**	0.72	0.26
20	0.30	0.24	**60**	0.64	0.62	**100**	0.72	0.68
21	0.21	0.18	**61**	0.52	0.28	**101**	0.24	0.23
22	0.81	0.61	**62**	0.79	0.68	**102**	1.05	0.31
23	0.56	0.30	**63**	0.53	0.18	**103**	0.11	0.10
24	0.32	0.25	**64**	0.55	0.21	**104**	0.30	0.19
25	0.54	0.45	**65**	0.46	0.40	**105**	1.03	0.69
26	0.32	0.23	**66**	0.46	0.43	**106**	0.69	0.45
27	0.42	0.22	**67**	0.41	0.25	**107**	0.45	0.25
28	0.74	0.32	**68**	0.84	0.54	**108**	0.83	0.60
29	0.82	0.12	**69**	0.63	0.41	**109**	1.01	0.43
30	0.59	0.44	**70**	0.69	0.50	**110**	1.02	0.60
31	0.64	0.32	**71**	1.15	0.57	**111**	0.67	0.34
32	0.60	0.48	**72**	0.43	0.39	**112**	0.45	0.29
33	0.60	0.52	**73**	0.72	0.26	**113**	0.42	0.35
34	1.27	0.43	**74**	0.60	0.51	**114**	0.27	0.22
35	1.18	1.15	**75**	0.43	0.43	**115**	0.38	0.21
36	0.81	0.59	**76**	0.91	0.28	**116**	0.66	0.67
37	0.64	0.36	**77**	0.47	0.38	**117**	0.30	0.24
38	0.39	0.15	**78**	0.72	0.28	**118**	0.90	0.78
39	0.22	0.17	**79**	0.60	0.40	**119**	0.54	0.45
40	1.08	0.75	**80**	0.31	0.31	**120**	0.35	0.33

Note: D.S. No. means Data Set Number

APPENDIX-3

"After 1900 We Begin To See Identifiable Statisticians Developing Such Techniques into Unified logic of Empirical Science That Goes for Beyond Its Components Parts." - Stephen J. Stigler

APPENDIX-3

FORWARD STEPWISE RIDGE REGRESSION MODEL

Forward stepwise ridge regression model results of the Japanese boys and girls are available to the researcher. Results as summary of them are shown below.

For boys

Dependent variable: Predicted final stature (*PFS*)

Table 1 Summary of forward stepwise ridge regression model for boys using ridge parameter: 0.0037, computed from modified HKB estimator

Step	Variable	Coefficient (*p*-value)	Standard error	F–to inter /remove (*p*-value)	DW-Value	R^2 (Adj. R^2)
1	*MS*	0.995856 (0.000001)	0.005087	38331.41 (0.000001)	2.0182	0.99535191 (0.99532594)
2	*MS*	0.511742 (0.000001)	0.038311	36396.53	2.0109	0.99756068
	S_9	0.486382 (0.000001)	0.038311	(0.000001)		(0.99753327)
3	*MS*	0.333572 (0.000001)	0.038531	34039.71	2.0213	0.99926973
	S_9	0.342368 (0.000001)	0.036509			
	FS	0.322908 (0.000001)	0.037915	(0.000001)		(0.99824040)
4	*MS*	0.272977 (0.000001)	0.036743	30447.29	2.0253	0.99955697
	S_9	0.229365 (0.000001)	0.038503			
	FS	0.268998 (0.000001)	0.035898	(0.000001)		(0.99852417)
	S_{10}	0.227805 (0.000001)	0.038488			
5	*MS*	0.233303 (0.000001)	0.035701	27314.93	2.0929	0.99972029
	S_9	0.186539 (0.000002)	0.037474			(0.99868373)
	FS	0.234572 (0.000001)	0.034676	(0.000001)		
	S_{10}	0.186001 (0.000002)	0.037409			
	S_3	0.158977 (0.000005)	0.033639			
6	*MS*	0.213970 (0.000001)	0.034889	24391.37	2.0918	0.99981247
	S_9	0.152222 (0.000071)	0.037388			(0.99877152)
	FS	0.214886 (0.000001)	0.033925	(0.000001)		
	S_{10}	0.148911 (0.000106)	0.037523			
	S_3	0.138225 (0.000044)	0.032985			
	S_{13}	0.131294 (0.000316)	0.035726			

Continued…

Step	Variable	Coefficient (*p*-value)	Standard error	F–to inter /remove (*p*-value)	DW-Value	R^2 (Adj. R^2)
7	*MS*	0.202341 (0.000001)	0.034292	21911.47	2.0823	0.99987336
	S_9	0.122862 (0.001371)	0.037763			
	FS	0.203117 (0.000001)	0.033362	(0.000001)		(0.99882777)
	S_{10}	0.122291 (0.001405)	0.037673			
	S_3	0.123091 (0.000219)	0.032599			
	S_{13}	0.111863 (0.001902)	0.035472			
	S_8	0.114002 (0.002585)	0.037283			
8	*MS*	0.193987 (0.000001)	0.033904	19795.92	2.1138	0.99991510
	S_9	0.107795 (0.013214)	0.037623	(0.000001)		(0.99886464)
	FS	0.195045 (0.000001)	0.032983			
	S_{10}	0.108592 (0.016611)	0.037457			
	S_3	0.102781 (0.008635)	0.033039			
	S_{13}	0.101755 (0.015176)	0.035130			
	S_8	0.097103 (0.006526)	0.037276			
	S_5	0.092558(0.027024)	0.035981			
9	*MS*	0.187364 (0.000001)	0.033699	17963.26	2.1089	0.99994340
	S_9	0.094847 (0.000001)	0.037726			(0.99888779)
	FS	0.189917 (0.000001)	0.032733	(0.000001)		
	S_{10}	0.093710 (0.000001)	0.037719			
	S_3	0.097045 (0.000001)	0.032810			
	S_{13}	0.087077 (0.000001)	0.035440			
	S_8	0.085253 (0.000001)	0.037307			
	S_5	0.084463 (0.000001)	0.035812			
	S_{11}	0.079968 (0.000001)	0.037361			

Table 2 Table for observed, predicted, residual, standard predicted, standard residual, standard error of predicted and Cook's distance values of predicted final stature of boys

Sl. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
01	170.6221	172.0385	-1.4164	-0.0826	-0.1953	0.6186	0.000094
02	171.5799	173.0780	-1.4981	0.0632	-0.2065	0.6845	0.000129
03	173.4181	172.6632	0.7549	0.0050	0.1041	0.9540	0.000065
04	182.9912	177.6363	5.3549	0.7024	0.7382	0.7061	0.001754
05	175.1818	174.2058	0.9760	0.2213	0.1346	0.6382	0.000047
06	170.7897	178.1026	-7.3130	0.7678	-1.0082	0.6394	0.002674
07	170.9299	172.4188	-1.4889	-0.0293	-0.2053	0.6033	0.000099
08	173.8874	168.3185	5.5690	-0.6043	0.7677	0.5401	0.001102
09	166.2960	164.3938	1.9022	-1.1546	0.2622	0.6019	0.000160
10	170.1347	170.8832	-0.7486	-0.2446	-0.1032	0.7758	0.000042
11	176.7793	179.3159	-2.5365	0.9379	-0.3497	0.8253	0.000542
12	175.2162	172.3364	2.8798	-0.0408	0.3970	0.6424	0.000419
13	171.6245	173.7799	-2.1554	0.1616	-0.2971	0.5903	0.000197
14	173.9456	175.0527	-1.1071	0.3401	-0.1526	0.5721	0.000049
15	180.3147	180.1309	0.1838	1.0522	0.0253	1.2105	0.000006

Continued…

D.S. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
16	175.7178	169.5234	6.1944	-0.4353	0.8540	0.6417	0.001932
17	166.7038	171.4104	-4.7066	-0.1707	-0.6489	0.5773	0.000900
18	170.3322	170.2104	0.1218	-0.3390	0.0168	0.6568	0.000001
19	168.5970	169.7184	-1.1214	-0.4080	-0.1546	0.5562	0.000047
20	168.8994	168.6840	0.2154	-0.5530	0.0297	0.6992	0.000003
21	177.2901	174.8690	2.4212	0.3143	0.3338	0.6630	0.000316
22	166.6451	172.0652	-5.4202	-0.0789	-0.7472	0.6262	0.001408
23	169.5538	169.1504	0.4034	-0.4876	0.0556	0.5354	0.000006
24	169.4539	172.1274	-2.6736	-0.0701	-0.3686	0.6565	0.000377
25	170.0412	172.3862	-2.3450	-0.0339	-0.3233	0.6360	0.000272
26	170.5650	173.6836	-3.1186	0.1481	-0.4299	0.6274	0.000468
27	169.0415	165.8468	3.1947	-0.9509	0.4404	0.5271	0.000345
28	171.2440	170.9932	0.2507	-0.2292	0.0346	0.6266	0.000003
29	173.9174	172.8338	1.0836	0.0289	0.1494	0.7467	0.000081
30	175.9431	169.9542	5.9889	-0.3749	0.8256	0.6260	0.001718
31	177.8915	177.3960	0.4955	0.6687	0.0683	0.5689	0.000010
32	167.4648	174.8593	-7.3945	0.3130	-1.0194	0.5667	0.002141
33	176.6420	172.7159	3.9261	0.0124	0.5413	0.5898	0.000654
34	172.2065	178.3625	-6.1561	0.8042	-0.8487	0.6203	0.001782
35	168.4455	174.7607	-6.3152	0.2991	-0.8706	0.6328	0.001953
36	173.4468	169.3221	4.1247	-0.4635	0.5686	0.5696	0.000673
37	168.2408	168.5999	-0.3591	-0.5648	-0.0495	0.5374	0.000005
38	175.6666	173.2984	2.3682	0.0941	0.3265	0.7858	0.000427
39	179.3212	175.2515	4.0697	0.3680	0.5611	0.5591	0.000631
40	165.5291	167.4789	-1.9498	-0.7220	-0.2688	0.5700	0.000151
41	173.3740	173.6956	-0.3216	0.1498	-0.0443	0.5525	0.000004
42	170.1054	174.4026	-4.2973	0.2489	-0.5924	0.5784	0.000753
43	173.7782	177.3176	-3.5394	0.6577	-0.4879	0.5559	0.000472
44	170.0710	166.7023	3.3687	-0.8309	0.4644	0.5842	0.000472
45	172.7364	168.1283	4.6082	-0.6309	0.6353	0.5716	0.000846
46	172.4249	173.5725	-1.1476	0.1325	-0.1582	0.6509	0.000068
47	164.6380	171.8912	-7.2532	-0.1033	-0.9999	0.5461	0.001911
48	171.7722	167.3839	4.3882	-0.7353	0.6050	0.6106	0.000877
49	167.0966	164.3523	2.7444	-1.1605	0.3783	0.5633	0.000291
50	173.1391	172.7129	0.4262	0.0120	0.0588	0.6325	0.000009
51	180.3872	177.5876	2.7995	0.6956	0.3860	0.6042	0.000349
52	172.3335	171.1696	1.1639	-0.2045	0.1605	0.5965	0.000059
53	180.2242	171.0364	9.1878	-0.2231	1.2666	0.6130	0.003875
54	170.0428	169.7345	0.3083	-0.4057	0.0425	0.5429	0.000003
55	171.9944	176.9578	-4.9634	0.6072	-0.6843	0.6776	0.001386
56	180.3848	171.6960	8.6889	-0.1306	1.1979	0.5704	0.002994
57	182.3506	176.5083	5.8423	0.5442	0.8054	0.6961	0.002029
58	168.8772	169.3022	-0.4249	-0.4663	-0.0586	0.6237	0.000009
59	165.9051	173.3373	-7.4322	0.0995	-1.0246	0.6083	0.002496
60	168.5783	168.5522	0.0260	-0.5715	0.0036	0.7301	0.000000
61	175.5057	170.5603	4.9454	-0.2899	0.6818	0.5597	0.000934
62	175.6034	174.6340	0.9694	0.2814	0.1336	0.6462	0.000048

Continued…

D.S. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
63	176.8999	176.1259	0.7741	0.4906	0.1067	0.6493	0.000031
64	164.1422	167.0685	-2.9263	-0.7795	-0.4034	0.6679	0.000468
65	184.6309	180.7033	3.9276	1.1325	0.5415	0.5679	0.000606
66	177.3452	178.1380	-0.7928	0.7727	-0.1093	0.6213	0.000030
67	175.6300	173.3417	2.2883	0.1002	0.3155	0.6114	0.000239
68	168.2848	172.1605	-3.8757	-0.0655	-0.5343	0.5769	0.000610
69	174.4799	180.7063	-6.2264	1.1329	-0.8584	0.7014	0.002340
70	178.0294	174.3458	3.6837	0.2410	0.5078	0.5684	0.000534
71	165.9994	161.1532	4.8462	-1.6091	0.6681	0.5738	0.000943
72	168.2601	170.9562	-2.6961	-0.2344	-0.3717	0.5779	0.000296
73	168.5236	172.5302	-4.0066	-0.0136	-0.5524	0.5718	0.000640
74	172.5603	173.7245	-1.1643	0.1538	-0.1605	0.5518	0.000050
75	175.2881	173.4261	1.8620	0.1120	0.2567	0.6162	0.000161
76	170.3066	174.0420	-3.7354	0.1984	-0.5150	0.5950	0.000603
77	180.3895	174.6245	5.7651	0.2800	0.7948	0.5811	0.001369
78	174.5264	176.6148	-2.0884	0.5591	-0.2879	0.6563	0.000230
79	168.6951	169.6977	-1.0026	-0.4109	-0.1382	0.5398	0.000036
80	169.0523	172.9066	-3.8543	0.0391	-0.5314	0.6417	0.000748
81	176.7590	173.7547	3.0043	0.1581	0.4142	0.5962	0.000392
82	173.3855	170.5828	2.8027	-0.2867	0.3864	0.5374	0.000276
83	178.5322	179.8478	-1.3157	1.0125	-0.1814	0.6827	0.000099
84	162.9345	169.0976	-6.1631	-0.4950	-0.8496	0.5593	0.001448
85	177.1721	169.9594	7.2127	-0.3742	0.9944	0.6247	0.002481
86	183.2272	177.9313	5.2959	0.7438	0.7301	0.7000	0.001686
87	164.7118	172.4751	-7.7633	-0.0214	-1.0703	0.8935	0.005974
88	168.3295	171.3122	-2.9827	-0.1845	-0.4112	0.7676	0.000645
89	176.1231	173.2982	2.8249	0.0940	0.3894	0.7547	0.000559
90	175.1326	175.6146	-0.4819	0.4189	-0.0664	0.6808	0.000013
91	172.0134	170.1585	1.8548	-0.3462	0.2557	0.5872	0.000145
92	177.0107	175.7663	1.2443	0.4402	0.1715	0.6410	0.000078
93	185.0216	184.4193	0.6023	1.6536	0.0830	0.5845	0.000015
94	171.7139	169.0919	2.6220	-0.4958	0.3615	0.6856	0.000396
95	172.5972	172.6495	-0.0523	0.0031	-0.0072	0.5405	0.000000
96	178.1410	176.5445	1.5965	0.5493	0.2201	0.6051	0.000114
97	168.1752	167.8981	0.2771	-0.6632	0.0382	0.5469	0.000003
98	176.5982	171.5978	5.0004	-0.1444	0.6894	0.5487	0.000917
99	168.2614	167.9956	0.2659	-0.6495	0.0367	0.5499	0.000003
100	166.9395	168.9537	-2.0143	-0.5152	-0.2777	0.5515	0.000150
101	176.1231	173.2982	2.8249	0.0940	0.3894	0.7547	0.000559
102	177.3322	175.3855	1.9467	0.3868	0.2684	0.6205	0.000178
103	161.2900	169.0541	-7.7640	-0.5011	-1.0704	0.7484	0.004153
104	168.0406	168.9070	-0.8664	-0.5217	-0.1194	0.6702	0.000041
105	184.9472	181.3362	3.6110	1.2212	0.4978	0.9590	0.001496
106	176.6280	176.0430	0.5850	0.4790	0.0807	0.6022	0.000015
107	175.1326	175.6146	-0.4819	0.4189	-0.0664	0.6808	0.000013
108	174.8539	176.5542	-1.7003	0.5507	-0.2344	0.5962	0.000125
109	170.5492	167.5953	2.9538	-0.7057	0.4072	0.6804	0.000495

Continued…

D.S. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
110	176.9766	176.4632	0.5135	0.5379	0.0708	0.9441	0.000029
111	171.1965	174.9355	-3.7390	0.3237	-0.5155	0.6469	0.000716
112	185.1041	179.7898	5.3143	1.0044	0.7326	0.7270	0.001834
113	174.7571	172.3408	2.4163	-0.0402	0.3331	0.5414	0.000208
114	168.2614	167.9956	0.2659	-0.6495	0.0367	0.5499	0.000003
115	161.1550	172.6561	-11.5010	0.0040	-1.5855	0.5675	0.005192
116	168.1752	167.8981	0.2771	-0.6632	0.0382	0.5469	0.000003
117	175.9952	171.4538	4.5414	-0.1646	0.6261	0.7313	0.001355
118	163.6482	172.0163	-8.3681	-0.0857	-1.1536	0.7890	0.005375
119	168.1752	171.9590	-3.7838	-0.0938	-0.5216	0.5452	0.000518
120	177.0218	172.0317	4.9901	-0.0836	0.6879	0.6514	0.001293
121	181.3418	177.6302	3.7116	0.7015	0.5117	0.7598	0.000979
122	179.9362	173.6558	6.2804	0.1442	0.8658	0.5516	0.001462
123	173.0444	166.2022	6.8422	-0.9010	0.9433	0.5224	0.001554
124	170.1188	172.2759	-2.1571	-0.0493	-0.2974	0.5431	0.000167
125	170.8584	168.5219	2.3365	-0.5757	0.3221	0.5312	0.000187
126	173.6972	173.0134	0.6839	0.0541	0.0943	0.6258	0.000022
127	162.9345	169.0976	-6.1631	-0.4950	-0.8496	0.5593	0.001448
128	178.6312	179.8478	-1.2166	1.0125	-0.1677	0.6827	0.000085
129	173.3855	170.5828	2.8027	-0.2867	0.3864	0.5374	0.000276
130	176.1629	174.2264	1.9365	0.2242	0.2670	0.5855	0.000157
131	169.0283	172.8987	-3.8704	0.0380	-0.5336	0.6420	0.000755
132	169.4439	169.6942	-0.2503	-0.4113	-0.0345	0.5397	0.000002
133	180.4058	174.6203	5.7855	0.2795	0.7976	0.5810	0.001378
134	170.1844	174.0749	-3.8905	0.2030	-0.5364	0.5950	0.000654
135	169.2182	172.5046	-3.2863	-0.0172	-0.4531	0.5726	0.000432
136	172.5603	173.7245	-1.1643	0.1538	-0.1605	0.5518	0.000050
137	175.2881	173.4261	1.8620	0.1120	0.2567	0.6162	0.000161
138	172.7382	175.2026	-2.4644	0.3611	-0.3397	0.5623	0.000234
139	175.9935	176.1210	-0.1274	0.4899	-0.0176	0.5722	0.000001
140	170.4385	170.3448	0.0937	-0.3201	0.0129	0.6641	0.000000
141	168.3295	171.3122	-2.9827	-0.1845	-0.4112	0.7676	0.000645
142	168.4455	174.7607	-6.3152	0.2991	-0.8706	0.6328	0.001953
143	165.1618	164.4240	0.7378	-1.1504	0.1017	0.5650	0.000021
144	174.4021	169.2482	5.1539	-0.4739	0.7105	0.5708	0.001055
145	172.2065	178.3625	-6.1561	0.8042	-0.8487	0.6203	0.001782
146	168.2493	168.6016	-0.3523	-0.5646	-0.0486	0.5374	0.000004
147	172.0821	173.2016	-1.1195	0.0805	-0.1543	0.7805	0.000094
148	170.2173	166.7266	3.4907	-0.8275	0.4812	0.5832	0.000505
149	170.6790	168.3464	2.3325	-0.6003	0.3216	0.5702	0.000216
150	172.3417	173.5771	-1.2354	0.1332	-0.1703	0.6510	0.000079
151	165.3860	171.7516	-6.3656	-0.1228	-0.8776	0.5485	0.001485
152	173.4419	172.6327	0.8092	0.0007	0.1116	0.6365	0.000032
153	169.5003	168.9423	0.5579	-0.5168	0.0769	0.8684	0.000029
154	175.1901	175.5192	-0.3291	0.4055	-0.0454	0.6600	0.000006
155	169.1957	174.3308	-5.1350	0.2389	-0.7079	0.9564	0.003008
156	172.9921	170.9995	1.9926	-0.2283	0.2747	0.6042	0.000177

Continued…

D.S. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
157	173.5821	174.7967	-1.2146	0.3042	-0.1674	0.5675	0.000058
158	178.1356	174.7861	3.3496	0.3027	0.4618	0.5805	0.000461
159	163.4866	168.2588	-4.7722	-0.6126	-0.6579	0.8473	0.002023
160	174.4793	175.1920	-0.7127	0.3596	-0.0983	0.6747	0.000028
161	170.0847	177.0656	-6.9809	0.6224	-0.9624	0.6147	0.002249
162	172.0748	168.1184	3.9564	-0.6323	0.5454	0.6896	0.000913
163	178.3780	171.9695	6.4085	-0.0923	0.8835	0.7315	0.002701
164	182.3472	176.5086	5.8385	0.5443	0.8049	0.6961	0.002026
165	168.8772	169.3022	-0.4249	-0.4663	-0.0586	0.6237	0.000009
166	180.8853	171.2596	9.6257	-0.1918	1.3270	0.6815	0.005274
167	178.4214	171.4324	6.9890	-0.1676	0.9635	0.5783	0.001992
168	173.3915	176.6237	-3.2321	0.5604	-0.4456	0.6969	0.000622
169	170.4899	169.7062	0.7837	-0.4097	0.1080	0.5434	0.000022
170	180.3727	171.0228	9.3498	-0.2250	1.2890	0.6132	0.004015
171	182.6243	177.3921	5.2322	0.6681	0.7213	0.6116	0.001251
172	172.5089	173.1368	-0.6280	0.0714	-0.0866	0.6914	0.000023
173	171.3062	172.1293	-0.8231	-0.0699	-0.1135	0.5779	0.000028
174	175.6300	173.3417	2.2883	0.1002	0.3155	0.6114	0.000239
175	168.6143	168.5389	0.0754	-0.5734	0.0104	0.7301	0.000000
176	169.0956	166.1521	2.9435	-0.9081	0.4058	0.6465	0.000443
177	164.1421	167.0685	-2.9264	-0.7795	-0.4034	0.6679	0.000468
178	183.2663	180.3776	2.8887	1.0868	0.3982	0.5706	0.000331
179	177.2191	178.1409	-0.9218	0.7732	-0.1271	0.6213	0.000040
180	175.3868	174.5981	0.7886	0.2763	0.1087	0.6442	0.000032

Note: D.S. No. means Data Set Number

For girls

Dependent variable: Predicted final stature (*PFS*)

Table 3 Summary of forward stepwise ridge regression model for girls using ridge parameter: ridge parameter: 0.0010, computed from modified HKB estimator

Step	Variable	Coefficient (*p*-value)	Standard error	F–to inter /remove (*p*-value)	DW-Value	R^2 (Adj. R^2)
1	S_{13}	0.998669 (0.000001)	0.003762	70478.91 (0.000001)	1.8534	0.99831440 (0.99830023)
2	S_{13}	0.559608 (0.000001)	0.048474	59367.00	1.7921	0.99900717
	MS	0.439848 (0.000001)	0.048474	(0.000001)		(0.99899034)
3	S_{13}	0.387948 (0.000001)	0.051731	50900.09	1.6557	0.99973438
	MS	0.316876 (0.000001)	0.047571	(0.000001)		(0.99921475)
	FS	0.294889 (0.000001)	0.050045			

Continued…

Step	Variable	Coefficient (*p*-value)	Standard error	F–to inter /remove (*p*-value)	DW-Value	R^2 (Adj. R^2)
4	S_{13}	0.329047 (0.000001)	0.052724	41447.49	1.5846	0.99973581
	MS	0.286799 (0.000001)	0.046546	(0.000001)		(0.99927670)
	FS	0.239452 (0.000001)	0.050850			
	S_2	0.144564 (0.001204)	0.043546			
5	S_{13}	0.269009 (0.000009)	0.058027	34375.80	1.5213	0.99975137
	MS	0.278478 (0.000001)	0.045858	(0.000001)		(0.99930230)
	FS	0.220903(0.000028)	0.050593			
	S_2	0.109648 (0.017303)	0.045399			
	S_9	0.121874 (0.023680)	0.053156			

Table 4 Table for observed, predicted, residual, standard predicted, standard residual, standard error of predicted and Cook's distance values of predicted final stature of girls

D.S. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
01	157.6700	154.5338	3.1362	-0.9476	0.7038	0.6917	0.004176
02	164.2200	163.0973	1.1227	0.9119	0.2519	0.4822	0.000254
03	157.6300	158.1579	-0.5279	-0.1607	-0.1185	0.5665	0.000078
04	160.2800	154.4489	5.8311	-0.9661	1.3085	0.7009	0.014843
05	163.8000	162.3004	1.4996	0.7389	0.3365	1.3067	0.003885
06	158.7600	154.8400	3.9200	-0.8812	0.8797	0.5278	0.003721
07	168.3700	166.6675	1.7025	1.6872	0.3820	0.5158	0.000670
08	164.1500	163.7488	0.4012	1.0534	0.0900	0.7186	0.000074
09	150.3400	153.7147	-3.3747	-1.1255	-0.7573	0.5886	0.003454
10	159.8200	161.6613	-1.8413	0.6001	-0.4132	0.4517	0.000597
11	160.3800	160.9394	-0.5594	0.4433	-0.1255	0.4791	0.000062
12	156.7700	159.3170	-2.5470	0.0910	-0.5716	0.4681	0.001228
13	154.4900	155.7193	-1.2293	-0.6902	-0.2759	0.4350	0.000246
14	157.9600	158.3664	-0.4064	-0.1154	-0.0912	0.5636	0.000046
15	157.4300	160.2552	-2.8252	0.2947	-0.6340	0.6122	0.002627
16	162.3100	160.8629	1.4471	0.4267	0.3247	0.4565	0.000377
17	157.4300	160.2552	-2.8252	0.2947	-0.6340	0.6122	0.002627
18	157.9600	158.3664	-0.4064	-0.1154	-0.0912	0.5636	0.000046
19	161.2400	160.8669	0.3731	0.4276	0.0837	0.4565	0.000025
20	160.6000	159.8486	0.7514	0.2065	0.1686	0.8332	0.000356
21	155.3700	154.6643	0.7057	-0.9193	0.1584	0.4748	0.000097
22	157.3400	155.2941	2.0459	-0.7826	0.4591	0.5933	0.001291
23	161.3600	157.5810	3.7790	-0.2860	0.8480	0.6013	0.004527
24	169.5600	161.2427	8.3173	0.5092	1.8664	0.5002	0.015005
25	166.0200	159.1167	6.9033	0.0475	1.5491	0.4121	0.006960
26	150.5900	154.5151	-3.9251	-0.9517	-0.8808	0.4965	0.003292
27	165.8600	160.1368	5.7232	0.2690	1.2843	0.4762	0.006425
28	158.0500	151.4489	6.6011	-1.6175	1.4813	0.4884	0.009002

Continued…

D.S. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
29	162.6300	161.3171	1.3129	0.5253	0.2946	0.5069	0.000384
30	161.1600	163.5680	-2.4080	1.0141	-0.5404	0.7080	0.002586
31	157.2500	158.8983	-1.6483	0.0001	-0.3699	0.4083	0.000389
32	162.1400	159.1924	2.9477	0.0640	0.6615	0.4865	0.001781
33	150.4200	150.0670	0.3530	-1.9176	0.0792	0.5887	0.000038
34	165.0600	160.2070	4.8530	0.2843	1.0890	0.4327	0.003798
35	156.7800	153.0070	3.7730	-1.2792	0.8467	0.8456	0.009259
36	155.7600	158.3503	-2.5903	-0.1189	-0.5813	0.4658	0.001258
37	165.9700	160.7882	5.1818	0.4105	1.1628	0.5754	0.007771
38	162.2200	161.1494	1.0706	0.4889	0.2403	0.4559	0.000206
39	162.6000	165.8806	-3.2806	1.5163	-0.7362	0.4565	0.001936
40	165.2900	160.4920	4.7980	0.3462	1.0767	0.4300	0.003665
41	165.3600	166.1404	-0.7804	1.5727	-0.1751	0.5055	0.000135
42	148.2100	151.3743	-3.1643	-1.6337	-0.7101	0.4925	0.002104
43	164.5100	161.9415	2.5685	0.6609	0.5764	0.4366	0.001084
44	157.9900	155.8874	2.1026	-0.6537	0.4718	0.5918	0.001356
45	153.5300	162.6664	-9.1364	0.8183	-2.0503	0.8423	0.053843
46	152.5900	156.5272	-3.9372	-0.5148	-0.8835	0.4059	0.002195
47	158.7900	156.7536	2.0364	-0.4656	0.4570	0.5771	0.001208
48	157.8200	160.9336	-3.1136	0.4421	-0.6987	0.5601	0.002654
49	166.2300	163.5931	2.6369	1.0196	0.5917	0.5404	0.001768
50	162.7300	161.3525	1.3775	0.5330	0.3091	0.6122	0.000624
51	156.1400	160.9056	-4.7656	0.4360	-1.0694	0.4433	0.003848
52	153.9400	160.3523	-6.4123	0.3158	-1.4390	0.5286	0.009992
53	161.5000	156.7614	4.7386	-0.4639	1.0634	0.4224	0.003448
54	154.5400	147.0753	7.4647	-2.5673	1.6751	0.5033	0.012244
55	160.5000	159.3877	1.1123	0.1064	0.2496	0.4197	0.000187
56	166.2600	163.8279	2.4321	1.0706	0.5458	0.4907	0.001233
57	156.2000	157.2815	-1.0815	-0.3510	-0.2427	0.5930	0.000360
58	159.0300	160.4232	-1.3932	0.3312	-0.3127	0.4735	0.000376
59	154.4300	158.2350	-3.8050	-0.1440	-0.8539	0.4093	0.002085
60	151.7400	157.9194	-6.1794	-0.2125	-1.3867	0.4212	0.005831
61	155.4100	156.4140	-1.0040	-0.5394	-0.2253	0.4884	0.000208
62	159.3200	159.1768	0.1432	0.0606	0.0321	0.4305	0.000003
63	152.1400	156.5903	-4.4504	-0.5011	-0.9987	0.4724	0.003822
64	159.0500	160.8668	-1.8168	0.4275	-0.4077	1.1323	0.004088
65	160.5200	164.1458	-3.6258	1.1396	-0.8136	0.4586	0.002387
66	151.4700	154.0715	-2.6016	-1.0480	-0.5838	0.5586	0.001843
67	152.8400	157.6272	-4.7872	-0.2759	-1.0743	0.4475	0.003958
68	153.8800	156.6798	-2.7998	-0.4817	-0.6283	0.5293	0.001910
69	160.8100	161.3723	-0.5623	0.5373	-0.1262	0.5923	0.000097
70	153.9300	158.3092	-4.3792	-0.1278	-0.9827	0.5044	0.004232
71	162.1100	163.3742	-1.2642	0.9720	-0.2837	0.5453	0.000414
72	157.6600	157.8711	-0.2111	-0.2230	-0.0474	0.4079	0.000006
73	165.3200	166.3763	-1.0562	1.6239	-0.2370	0.4384	0.000185
74	154.3800	156.4169	-2.0369	-0.5387	-0.4571	0.5291	0.001010
75	165.8200	164.6733	1.1467	1.2541	0.2573	0.5747	0.000380
76	156.8800	157.3223	-0.4423	-0.3421	-0.0993	0.4659	0.000037

Continued…

D.S. No.	Observed Value	Predicted Value	Residual	Standard Predicted Value	Standard Residual	Standard Error of Predicted Value	Cook's Distance
77	159.8800	160.9253	-1.0453	0.4403	-0.2346	0.6096	0.000356
78	157.9200	157.4142	0.5058	-0.3222	0.1135	0.5855	0.000077
79	160.3000	158.7440	1.5560	-0.0334	0.3492	0.5525	0.000644
80	163.5200	161.9684	1.5516	0.6668	0.3482	0.6056	0.000775
81	161.0500	159.6740	1.3760	0.1685	0.3088	0.5848	0.000567
82	162.0100	164.5820	-2.5720	1.2343	-0.5772	0.5820	0.001960
83	163.4100	157.3518	6.0582	-0.3357	1.3595	0.6118	0.012061
84	161.6042	159.0135	2.5907	0.0251	0.5814	0.6191	0.002261
85	162.4200	162.5588	-0.1388	0.7950	-0.0312	0.6534	0.000007
86	158.7500	159.3093	-0.5593	0.0894	-0.1255	0.4336	0.000051
87	157.3400	155.4660	1.8740	-0.7452	0.4205	0.5182	0.000819
88	158.1600	161.1474	-2.9874	0.4885	-0.6704	0.4651	0.001668
89	163.4200	159.8714	3.5486	0.2114	0.7963	0.7685	0.006677
90	161.6379	160.2795	1.3584	0.3000	0.3048	0.4369	0.000304
91	151.4200	156.2678	-4.8478	-0.5711	-1.0879	0.4494	0.004094
92	148.9300	155.0689	-6.1389	-0.8315	-1.3776	0.5317	0.009268
93	163.6400	159.8000	3.8400	0.1959	0.8617	0.4736	0.002860
94	155.0800	158.2487	-3.1687	-0.1410	-0.7111	0.5255	0.002410
95	153.3000	157.0817	-3.7817	-0.3944	-0.8486	0.4794	0.002844
96	157.7500	154.0202	3.7298	-1.0592	0.8370	0.7336	0.006686
97	158.8200	154.6381	4.1819	-0.9250	0.9384	0.4588	0.003178
98	162.2900	160.4711	1.8189	0.3416	0.4082	0.6703	0.001315
99	162.5000	162.1436	0.3564	0.7048	0.0800	0.4385	0.000021
100	160.8300	158.7193	2.1108	-0.0388	0.4737	0.4877	0.000917
101	158.2900	161.6096	-3.3197	0.5889	-0.7449	0.5536	0.002945
102	158.0300	160.4656	-2.4356	0.3404	-0.5466	0.5242	0.001417
103	158.7700	153.2875	5.4825	-1.2183	1.2303	0.5169	0.006975
104	151.0700	154.6846	-3.6146	-0.9149	-0.8111	0.5829	0.003884
105	154.2500	156.7295	-2.4795	-0.4709	-0.5564	0.4089	0.000884
106	154.3700	154.4865	-0.1165	-0.9579	-0.0261	0.5185	0.000003
107	157.9400	163.2141	-5.2741	0.9373	-1.1835	0.6659	0.010907
108	159.5400	157.8173	1.7227	-0.2346	0.3866	0.4440	0.000505
109	154.1300	155.7142	-1.5842	-0.6913	-0.3555	0.5335	0.000621
110	160.0800	160.1757	-0.0957	0.2775	-0.0215	0.5327	0.000002
111	155.2600	157.7967	-2.5367	-0.2391	-0.5692	0.6346	0.002282
112	154.9200	153.2544	1.6656	-1.2255	0.3738	0.4251	0.000432
113	157.1200	158.7710	-1.6511	-0.0275	-0.3705	0.5475	0.000712
114	157.3100	157.7881	-0.4781	-0.2410	-0.1073	0.4707	0.000044
115	164.6000	167.5579	-2.9579	1.8805	-0.6638	0.4970	0.001873
116	156.9800	155.3224	1.6576	-0.7764	0.3720	0.5906	0.000839
117	159.4100	159.8567	-0.4467	0.2082	-0.1002	0.8339	0.000126
118	160.1400	158.3173	1.8227	-0.1261	0.4090	0.4276	0.000523
119	166.0200	159.1167	6.9033	0.0475	1.5491	0.4121	0.006960
120	164.6800	161.5016	3.1784	0.5654	0.7133	0.5200	0.002373

Note: D.S. No. means Data Set Number

End